

# Multi-Attribute Proportional Representation

Jérôme Lang  
Université Paris-Dauphine  
Paris, France

Piotr Skowron  
University of Warsaw  
Warsaw, Poland

## Abstract

We consider the following problem in which a given number of items has to be chosen from a predefined set. Each item is described by a vector of attributes and for each attribute there is a desired distribution that the selected set should have. We look for a set that fits as much as possible the desired distributions on all attributes. Examples of applications include choosing members of a representative committee, where candidates are described by attributes such as sex, age and profession, and where we look for a committee that for each attribute offers a certain representation, i.e., a single committee that contains a certain number of young and old people, certain number of men and women, certain number of people with different professions, etc. With a single attribute the problem collapses to the apportionment problem for party-list proportional representation systems (in such case the value of the single attribute would be a political affiliation of a candidate). We study the properties of the associated subset selection rules, as well as their computation complexity.

## 1 Introduction

A research department has to choose  $k$  members for a recruiting committee. A selected committee should be gender balanced, ideally containing 50% of male and 50% of female. Additionally, a committee should represent different research areas in certain proportions: ideally it should contain 55% of researchers specializing in area  $A$ , 25% of experts in area  $B$ , and 20% in area  $C$ . Another requirement is that the committee should contain 30% junior and 70% senior researchers, and finally, the repartition between local and external members should be kept in proportions 30% to 70 %. The pool of possible members is the following:

Name	Sex	Group	Age	Affiliation
Ann	$F$	$A$	$J$	$L$
Bob	$M$	$A$	$J$	$E$
Charlie	$M$	$A$	$S$	$L$
Donna	$F$	$B$	$S$	$E$
Ernest	$M$	$A$	$S$	$L$
George	$M$	$A$	$S$	$E$
Helena	$F$	$B$	$S$	$E$
John	$M$	$B$	$J$	$E$
Kevin	$M$	$C$	$J$	$E$
Laura	$F$	$C$	$J$	$L$

In the given example, if the department wants to select  $k = 3$  members, then it is easy to see that there exists no such committee that would ideally satisfy all the criteria. Nevertheless, some committees are better than others: intuitively we feel the sex ratio should be either equal to 2:1 or to 1:2, the area ratio should be equal to 2:1:0, the age ratio to 1:2, and the affiliation ratio to 1:2. Such relaxed criteria can be achieved by selecting Ann, Donna, and George. Now, let us consider the above example for the case when  $k = 4$ . In such case, the ideal sex ratio should be equal to 2:2, the research area ratio to 2:1:1, the age ratio to 1:3, and the

affiliation ratio to 1:3. It can be proved, however, that for  $k = 4$  there exists no committee satisfying such relaxed criteria. Intuitively, in such case the best committee is either {Ann, Charlie, Donna, George}, with two externals instead of three, or {Charles, Donna, George, Kevin}, with males being over-represented.

In this paper we formalize the intuition given in the above example and define what it means for a committee to be optimal. When looking for an appropriate definition we follow an axiomatic approach. First, we notice that our model generalizes the *apportionment* problem for proportional representation [2]. The central question of the apportionment problem is how to distribute parliament seats between political parties, given the numbers of votes casted for each party. Indeed, we can consider our multi-attribute problem, with the single attribute being a political affiliation of a candidate, and the desired distributions being the proportions of votes casted for different parties. In such case we can see that selecting a committee in our multi-attribute proportional representation system boils down to selecting a parliament according to some apportionment method.

There is a variety of apportionment methods studied in the literature [1]. In this paper we do not review these methods in detail (we refer the reader to the survey of Balinski and Young [2]), but we rather focus on a specific set of their properties that have been analyzed, namely *non-reversal*, *exactness* and *respect of quota*, *population monotonicity*, and *house monotonicity*. We define the analogs of these properties for the multi-attribute domain, and analyze our definition of an optimal committee for a multi-attribute domain with respect to these properties.

To emphasize the analogy between our model and the apportionment methods, we should provide some discussion on where the desired proportions for attributes come from. Typically, but not always, they come from *votes*. For instance, each voter might give her preferred value for each attribute, and the ideal proportions coincide with the observed frequencies. For instance, out of 20 voters, 10 would have voted for a male and 10 for a female, 13 for a young person and 7 for a senior one, etc. It is worth mentioning that the voters might cast approval ballots, that is for each attribute they might define a set of approved values rather than pointing out the single most preferred one. On the other hand, sometimes, instead of votes, there are “global” preferences on the composition of the committee, expressed directly by the group, imposed by law, or by other constraints that should be respected as much as possible independently of voter preferences.

The multi-attribute case, however, is also substantially different from the single-attribute one. In particular, multi-attribute proportional representation systems exhibit computational problems that do not appear in the single-attribute setting. Indeed, in the second part of our paper we show that finding an optimal committee is often NP-hard. However, we show that this challenge can be addressed by designing efficient approximation and fixed-parameter tractable algorithms.

After positioning our work with respect to related areas in Section 2, we present our model in Section 3. In Sections 4 and 5 we discuss relevant properties of methods for multi-attribute fair representation. In Section 6 we show that, although the computational of optimal committees is generally NP-hard, there exist good approximation and fixed-parameter tractable algorithms for finding them. In Section 7 we point to further research issues.

## 2 Related work

Our model is related to three distinct research areas:

**Voting on multi-attribute domains** (see the work of Lang and Xia [13] for a survey). There, the aim is to output a single winning combination of attributes (*e.g.*, in multiple referenda, a combination of binary values). Our model in case when  $k = 1$  can be viewed as a voting problem on a *constrained* multi-attribute domain (constrained because not all combinations are feasible).

**Multiwinner (or committee) elections.** In particular, our model is related to the problem of finding a *fully proportional representation* [6, 18]. There, the voters vote directly for candidates and do not consider attributes that characterize them. Thus, in this literature, the term “proportional representation” has a different meaning: these methods are ‘representative’ because each voter feels represented by some member of the elected committee. The computational aspects of full proportional and its extensions have raised a lot of

attention lately [21, 3, 7, 24, 17]. Our study of the properties of multi-attribute proportional representation is close in spirit to the work of Elkind et al. [10], who gives a normative study of multiwinner election rules. *Budgeted social choice* [16] is technically close to committee elections, but it has a different motivation: the aim is to make a collective choice about a set of objects to be consumed by the group (perhaps, subject to some constraints) rather than about the set of candidates to represent voters.

**Apportionment for party-list representation systems** (see the work of Balinski and Young [2] for a survey). As we already pointed out, the apportionment methods correspond to the restriction of our model to a single attribute (albeit with a different motivation). While voting on multi-attribute domains and multiwinner elections have lead to significant research effort in computational social choice, this is less the case for party-list representation systems. Ding and Lin [8] studied a game-theoretic model for a party-list proportional representation system under specific assumptions, and show that computing the Nash equilibria of the game is hard. Also related is the computation of bi-apportionment (assignment of seats to parties within regions), investigated in a few recent papers [22, 23, 14].

**Constrained approval voting** (CAP) [4, 20] is probably the closest work to our setting (MAPR). In CAP there are also multiple attributes, candidates are represented by tuples of attribute values, there is a target composition of the committee and we try to find a committee close to this target. However, there are also substantial differences between MAPR and CAP. First, in CAP, the target composition of the committee, exogenously defined, consists of a target number of seats *for each combination of attributes* (called a cell), that is, for each  $\vec{z} \in D_1 \times \dots \times D_p$ , we have a value  $s(\vec{z})$ ; while in MAPR we have a smaller input consisting of a target number for each value of each attribute. Note that the input in CAP is exponentially large in the number of attributes, which makes it infeasible in practice as soon as this number exceeds a few units (probably CAP was designed only for very small numbers of attributes, such as 2 or 3). Second, in CAP, the selection criterion of an optimal committee is made in two consecutive steps: first a set of *admissible committees* is defined, and the choice between these admissible committees is made by using approval ballots, and the chosen committee is the admissible committee maximizing the sum, over all voters, of the number of candidates approved (there is no loss function to minimize as in MAPR). A simple translation of CAP into an integer linear programming problem is given in [20, 25].

### 3 The model

Let  $X = \{X_1, \dots, X_p\}$  be a set of  $p$  attributes, each with a finite domain  $D_i = \{x_i^1, \dots, x_i^{q_i}\}$ . We say that  $X_i$  is binary if  $|D_i| = 2$ . We let  $D = D_1 \times \dots \times D_p$ . Let  $C = \{c_1, \dots, c_m\}$  be a set of candidates, also referred to as the *candidate database*. Each candidate  $c_i$  is represented as a vector of attribute values  $(X_1(c_i), \dots, X_p(c_i)) \in D$ .<sup>1</sup>

For each  $i \leq p$ , by  $\pi_i$  we denote a *target distribution*  $\pi_i = (\pi_i^1, \dots, \pi_i^{q_i})$  with  $\sum_{j=1}^{q_i} \pi_i^j = 1$ . We set  $\pi = (\pi_1, \dots, \pi_p)$ . Typically,  $n$  voters have casted a ballot expressing their preferred value on every attribute  $X_i$ , and  $\pi_i^j$  is the fraction of voters who have  $x_i^j$  as their preferred value for  $X_i$ , but the results presented in the paper are independent from where the values  $\pi_i^j$  come from (see the discussion in the Introduction).

The goal is to select a committee<sup>2</sup> of  $k \in \{1, \dots, m\}$  candidates (or items) such that the distribution of attribute values is as close as possible to  $\pi$ . Formally, let  $S_k(C)$  denote the set of all subsets of  $C$  of cardinality  $k$ . Given  $A \in S_k(C)$ , the *representation vector* for  $A$  is defined as  $r(A) = (r_1(A), \dots, r_p(A))$ , where  $r_i(A) = (r_i^j(A) | 1 \leq j \leq q_i)$  for each  $i = 1, \dots, p$ , and  $r_i^j(A) = \frac{|\{c \in A : X_i(c) = x_i^j\}|}{k}$ .

**Definition 1** A committee  $A \in S_k(C)$  is perfect for  $\pi$  if  $r_i(A) = \pi_i$  for all  $i$ .

<sup>1</sup>By writing  $X_j(c_i)$ , we slightly abuse notation, that is, we consider  $X_j$  both as an attribute name and as a function that maps any candidate to an attribute value; this will not lead to any ambiguity.

<sup>2</sup>We will stick to the terminology “committee” although the meaning of subsets of candidates has sometimes nothing to do with the election of a committee.

Thus, a perfect committee matches exactly the target distribution. Clearly, there is no perfect committee if for some  $i, j$ ,  $\pi_i^j$  is not an integer multiplicity of  $\frac{1}{k}$ . In some of our results we will focus on target distributions such that for each  $i, j$  the value  $k\pi_i^j$  is an integer. We will refer to such target distributions as to *natural* distributions.

We define metrics measuring how well a committee fits a target distribution, called *loss functions*.

**Definition 2** A loss function  $f$  maps  $\pi$  and  $r$  to  $f(\pi, r(A)) \in \mathbb{R}$ , and satisfies  $f(\pi, r(A)) = 0$  if and only if  $\pi = r$ .

There are a number of loss functions that can be considered. As often, the most classical loss functions use  $L^p$  norms, with the most classical examples of  $L^1$ ,  $L^2$ , and  $L^\infty$ . We focus on two representative  $L^p$  norms,  $L^1$ , and  $L^\infty$ , but we believe that other choices are also justified and may lead to interesting variants of our model. Consequently, we consider the following loss functions:

- $\|\cdot\|_1 : \|\pi, r(A)\|_1 = \sum_{i,j} |r_i^j(A) - \pi_i^j|$ .
- $\|\cdot\|_{1,\max} : \|\pi, r(A)\|_{1,\max} = \sum_i \max_j |r_i^j(A) - \pi_i^j|$ .
- $\|\cdot\|_{\max} : \|\pi, r(A)\|_{\max} = \max_{i,j} |\pi_i^j - r_i^j(A)|$ .

Now, we are ready to formally define the central problem addressed in the paper.

**Definition 3 (OPTIMAL REPRESENTATION)** Given  $X$ ,  $C$ ,  $\pi$ ,  $k$ , and a loss function  $f$ , find a committee  $A \in S_k(C)$  minimizing  $f(\pi, r(A))$ .

**Example 1** For the example of the Introduction, we have  $X = \{\text{sex, group, age, affiliation}\}$ ,  $D = \{F, M\} \times \{A, B, C\} \times \{J, S\} \times \{L, E\}$ , and  $X_1(\text{Ann}) = F$ ,  $X_1(\text{Bob}) = M$  etc.  $\{\text{Charlie, Donna, George, Kevin}\}$  is optimal for  $\|\cdot\|_1$ , with  $\|\pi, r(A)\|_1 = 0.5 + 0.1 + 0.1 + 0.1 = 0.8$ , and for  $\|\cdot\|_{1,\max}$ , with  $\|\pi, r(A)\|_{1,\max} = 0.4$ , but not for  $\|\cdot\|_{\max}$ .  $\{\text{Ann, Charlie, Donna, George}\}$  is optimal for  $\|\cdot\|_{\max}$ , with  $\|\pi, r(A)\|_{\max} = \max(0, 0.2, 0.05, 0.2) = 0.2$ , but not for the other criteria.

## 4 The single-attribute case

In this section we focus on the single-attribute case ( $p = 1$ ). Without loss of generality, let us assume that the single attribute be party affiliation. Further, let us for a moment assume that for each value  $x_1^j$  there are at least  $k$  candidates with value  $x_1^j$  (this is typically the case in party-list elections). Then finding the optimal committee comes down to apportionment problem for party-list elections, where a fractional distribution  $\pi_1$  has to be “rounded up” to an integer-valued distribution  $r_1$  such that  $\sum_j r_1^j = k$ .

There are two main families of apportionment methods: *largest remainders* and *highest average* methods [2]. We shall not discuss highest average methods here, because they are weakly relevant to our model. For largest remainders methods, a *quota*  $q$  is computed as a function of the number of seats  $k$  and the number of voters  $n$ . The number of votes for party  $i$  is  $n_i = n \cdot \pi_i$ . The most common choice of a quota is the *Hare quota*, defined as  $\frac{n}{k}$ ; the method based on the Hare quota is called the *Hamilton method*.<sup>3</sup> Our aim is to generalize the Hamilton method to multiattribute domains.

**Definition 4 (The largest remainder method.)** The largest remainder method with quota  $q$  is defined as follows:

- for all  $i$ ,  $s_i^* = \frac{n_i}{q}$  is the ideal number of seats for party  $i$ .
- each party  $i$  receives  $s_i = \lfloor s_i^* \rfloor$  seats; let  $t_i = s_i - s_i^*$  (called the remainder).

---

<sup>3</sup>Other common choices are the *Droop quota*  $1 + \frac{n}{1+k}$ , the *Hagenbach-Bischoff quota*  $\frac{n}{1+k}$  and the *Imperiali quota*  $\frac{n}{2+k}$ .

- the remaining  $k - \sum_i s_i$  seats are given to the  $k - \sum_i s_i$  parties with the highest remainders  $t_i$ .

Below we show that the largest remainder methods select a distribution  $(k_1, \dots, k_q)$  minimizing  $\max_{i=1, \dots, p} (s_i^* - k_i)$ , which in the case of Hamilton comes down to minimizing  $\max_{i=1, \dots, p} (\frac{n_i}{q} - k_i)$ . After defining  $\pi_1^i = \frac{n_i}{n}$  for all  $i$ , we obtain the result that explains that our problem, with any of the three variants of loss functions, generalizes the Hamilton apportionment method.

**Proposition 1** *When  $p = 1$  and assuming there are at least  $k$  items for each attribute, optimal subsets for  $\|\cdot\|_1$ ,  $\|\cdot\|_{1, \max}$  and  $\|\cdot\|_{\max}$  coincide, and correspond to the subsets given by the Hamilton apportionment method.*

*Proof.* Note that  $\|\cdot\|_{1, \max}$  and  $\|\cdot\|_{\max}$  are equivalent for  $p = 1$ . Recall that  $s_j^*$  denotes the target number of seats for party  $j$ . Let  $A$  be a committee of size  $k$  and let  $R^j(A) = k r^j(A)$  be the number of members of  $A$  that belong to party  $j$ . Since  $|R^j(A) - s_j^*| = k|r^j(A) - \pi^j|$ , we need to show that the following three assertions are equivalent:

1.  $A$  minimizes  $\sum_j |R^j(A) - s_j^*|$ .
2.  $A$  minimizes  $\max_j |R^j(A) - s_j^*|$ .
3.  $A$  is a Hamilton committee.

We first show  $1 \Rightarrow 3$ . Assume  $A$  is not a Hamilton committee: then there exists an attribute value (party) that receives strictly more or strictly less seats than it would receive according to the Hamilton method. Naturally, there must also exist an attribute that receives strictly less or strictly more seats, respectively. Formally, this means that there are two attribute values (parties), say 1 and 2, such that the target number of seats for parties 1 and 2 are  $s_1^* = p + \alpha_1$  and  $s_2^* = q + \alpha_2$ , with  $p, q$  integers and  $1 > \alpha_2 > \alpha_1 \geq 0$ , and such that either  $R^1(A) \geq p + 1$  and  $R^2(A) \leq q$ . We have  $\sum_j |R^j(A) - s_j^*| = \sum_{j \neq 1, 2} |R^j(A) - s_j^*| + |R^1(A) - s_1^*| + |R^2(A) - s_2^*| \geq \sum_{j \neq 1, 2} |R^j(A) - s_j^*| + (1 - \alpha_1) + \alpha_2$ . Consider the committee  $A'$  obtained from  $A$  by giving one less seat to 1 and one more to 2.

- If  $R^1(A) > p + 1$  then  $\sum_j |R^j(A) - s_j^*| - \sum_j |R^j(A') - s_j^*| = |R^1(A) - s_1^*| - |R^1(A') - s_1^*| + |R^2(A) - s_2^*| - |R^2(A') - s_2^*| \geq 1 + (1 - \alpha_2) - \alpha_2 > 0$ .
- If  $R^2(A) < q$  then similarly,  $\sum_j |R^j(A) - s_j^*| - \sum_j |R^j(A') - s_j^*| > 0$ .
- If  $R^1(A) = p + 1$  and  $R^2(A) = q$  then we have  $\sum_j |R^j(A) - s_j^*| = \sum_{j \neq 1, 2} |R^j(A) - s_j^*| + (1 - \alpha_1) + \alpha_2$  and  $\sum_j |R^j(A') - s_j^*| = \sum_{j \neq 1, 2} |R^j(A') - s_j^*| + (1 - \alpha_2) + \alpha_1$ , hence  $\sum_j |R^j(A) - s_j^*| - \sum_j |R^j(A') - s_j^*| = 2(\alpha_2 - \alpha_1) > 0$ .

In all three cases,  $A$  does not minimize  $\sum_j |R^j(A) - s_j^*|$  and is therefore not an optimal committee for  $\|\cdot\|_{1, \Sigma}$ .

We now show  $2 \Rightarrow 3$ . Call a party  $i$  *lucky* if  $R^i(A) > s_i^*$  and *unlucky* if  $R^i(A) < s_i^*$ . Then we have  $\max_i |R^i(A) - s_i^*| = \max(0, \max\{R^i(A) - s_i^* | i \text{ lucky}\}, \max\{s_i^* - R^i(A) | i \text{ unlucky}\})$ . Let, without loss of generality, 1 be the lucky party with the highest value (if there are several such parties, we take arbitrary one of them)  $R^1(A) - s_1^*$  and 2 be the unlucky party with the highest value  $s_2^* - R^2(A)$ . Assume  $A$  is not a Hamilton committee: then 2 had a higher remainder than 1 before 1 got her last seat, that is,  $R^2(A) - s_2^* > (R^1(A) - 1) - s_1^*$ . Let  $A'$  be the committee  $A'$  obtained from  $A$  by giving one less seat to 1 and one more to 2: then either  $A'$  is a Hamilton committee, or it is not, and in this case we repeat the operation until we get a Hamilton committee  $A^*$ . Because  $\max_j |R^j(A^*) - s_j^*| < \max_j |R^j(A) - s_j^*|$ ,  $A$  is not an optimal committee for  $\|\cdot\|_{\max}$ .

It remains to be shown that if  $A$  is a Hamilton committee then it is both optimal for  $\|\cdot\|_{1, \max}$  and  $\|\cdot\|_{\max}$ . If there is a unique Hamilton-optimal committee then this follows immediately from  $1 \Rightarrow 3$  and  $2 \Rightarrow 3$ . Assume there are several Hamilton-optimal committees  $A_1, \dots, A_q$ . Then there are  $q$  parties, w.l.o.g.,

$1, \dots, q$ , with equal remainders  $\alpha \in [0, 1)$ , that is,  $s_1^* = p_1 + \alpha, \dots, s_q^* = p_q + \alpha$ , and the Hamilton-optimal committees differ only in the choice if those of these  $q$  parties to give they give an extra seat. We easily check that for any two  $A, A'$  of these committees we have  $\|A\|_{1,\max} = \|A'\|_{1,\max}$  and  $\|A\|_{\max} = \|A'\|_{\max}$ .  $\square$

Therefore, our model can be seen as a generalization of the Hamilton apportionment method to more than attribute. Note that our model can easily extend other largest remainder methods, and our results would be easily adapted. Interestingly, when  $p \geq 2$ , our three criteria no longer coincide. However, for binary domains,  $\|\cdot\|_1$  and  $\|\cdot\|_{1,\max}$  coincide, since  $\sum_{j=1,2} |r_i^j(A) - \pi_i^j| = 2 \max_{j=1,2} |r_i^j(A) - \pi_i^j|$ .

### Proposition 2

1. For each  $p \geq 3$  and binary domains, optimal subsets for  $\|\cdot\|_1$  and  $\|\cdot\|_{\max}$  may be disjoint, even for  $k = 2$ .
2. For each  $p \geq 3$ , optimal subsets for  $\|\cdot\|_{\max}$  and  $\|\cdot\|_{1,\max}$  can be disjoint.
3. For each  $p \geq 2$ , if at least one attribute has 4 values, then optimal subsets for  $\|\cdot\|_1$  and  $\|\cdot\|_{1,\max}$  can be disjoint.
4. For  $p = 2$  and binary domains, optimal subsets for  $\|\cdot\|_1$  and  $\|\cdot\|_{\max}$  may differ.

*Proof.* We prove point 1 for  $p = 3$  (the proof extends easily to  $p > 3$  by adding attributes on which all items, and the target, agree). We have four candidates: two ( $A$  and  $B$ ) with attribute vectors  $(x_1^2, x_2^1, x_3^1)$ , and two ( $C$  and  $D$ ) with  $(x_1^1, x_2^2, x_3^2)$ . The target distribution is  $\pi_i^1 = 0$  and  $\pi_i^2 = 1$  for  $i \in \{1, 2, 3\}$ . The  $\|\cdot\|_{\max}$ -optimal committees are  $\{A, C\}, \{A, D\}, \{B, C\}$  and  $\{B, D\}$ . The  $\|\cdot\|_1$ -optimal committee is  $\{C, D\}$ .

For Point 2: because optimal subsets for  $\|\cdot\|_1$  and  $\|\cdot\|_{1,\max}$  coincide for binary domains, Point 1 implies that optimal subsets for  $\|\cdot\|_{\max}$  and  $\|\cdot\|_{1,\max}$  can be disjoint. The counterexample extends easily to non-binary domains.

For Point 3: Let there be two attributes  $X_1$  with values  $x_1^1, x_1^2, x_1^3, x_1^4$  and  $X_2$  with values  $x_2^1, x_2^2$ ; four candidates:  $A$  with value vector  $(x_1^1, x_2^2)$ ,  $B$  with value vector  $(x_1^2, x_2^2)$ ,  $C$  with value vector  $(x_1^3, x_2^1)$ , and  $D$  with value vector  $(x_1^4, x_2^1)$ ;  $k = 2$ ; and  $\pi = (0.5, 0.5, 0, 0)$  for  $X_1$  and  $(0.9, 0.1)$  for  $X_2$ . The optimal committees for  $\|\cdot\|_1$  are all pairs except  $\{C, D\}$  (with loss 1.8) while the optimal committee for  $\|\cdot\|_{1,\max}$  is  $\{C, D\}$  (with loss 0.6). Next, we show that  $\|\cdot\|_{\max}$  and  $\|\cdot\|_{1,\max}$  can be disjoint. The counterexample extends easily to more attributes and more values.

For Point 4, let  $k = 2$ , three candidates  $A, B$  and  $C$  with value vectors  $(x_1^1, x_2^1)$ ,  $(x_1^1, x_2^1)$  and  $(x_1^2, x_2^2)$ ; and  $\pi_1^1 = 1, \pi_1^2 = 0, \pi_2^1 = 0$ , and  $\pi_2^2 = 1$ .  $\{A, B\}, \{A, C\}$  and  $\{B, C\}$  are all  $\|\cdot\|_1$ -optimal, but only  $\{A, C\}$  and  $\{B, C\}$  are  $\|\cdot\|_{\max}$ -optimal.  $\square$

These negative results come from the constraints imposed by the candidate database, which prevent the selection on the different attributes to be done independently. In the example of the proof of point 1, for instance, since all items with the value  $x_2^1$  for  $X_2$  have value  $x_3^1$  for  $X_3$ , selecting  $q$  items with  $X_2 = x_2^1$  implies selecting  $q$  items with  $X_3 = x_3^1$ . However, if the database is sufficiently diverse so that no such constraints exist, the optimization can be done separately on each attribute. This is captured by the following notion.

**Definition 5** A candidate database  $C$  satisfy the Full Supply (FS) property with respect to  $k$  if for any  $\vec{x} \in D$  there are at least  $k$  candidates in  $C$  associated with value vector  $\vec{x}$ .

The candidate database of Example 1 does not satisfy FS, even for  $k = 1$ , because there is not a single candidate with group  $C$  and age  $S$ . If we ignore attributes *group* and *affiliation*, then we are left with 2 (resp., 3, 2, 3) candidates with value vector  $FJ$  (resp.  $MJ, FS, MS$ ): the reduced database satisfies FS for  $k \in \{1, 2\}$ .

**Proposition 3** Let  $(X, C, k)$  be an optimal committee selection problem. If  $C$  satisfies FS w.r.t.  $k$ , then the following statements are equivalent:

- $A$  is an optimal committee for  $\|\cdot\|_1$
- $A$  is an optimal committee for  $\|\cdot\|_{1,\max}$
- for any attribute  $X_i$ ,  $A$  is a Hamilton committee for the single-attribute problem  $(\{X_i\}, D^{\downarrow X_i}, \pi_i, k)$ , where  $D^{\downarrow X_i}$  is the projection of  $D$  on  $\{X_i\}$ .

Moreover, any  $\|\cdot\|_1$  (and  $\|\cdot\|_{1,\max}$ ) optimal committee is optimal for  $\|\cdot\|_{\max}$ . (The converse does not always hold.)

*Proof.* For each attribute  $X_i$  and value  $x_i^j \in D_i$ , let  $R_i^j$  be the number of seats with value  $x_i^j$  given by the Hamilton method for the single-attribute problem  $(\{X_i\}, D^{\downarrow X_i}, \pi_i, k)$ . For all  $j = 1, \dots, k$ , let  $t_i(j) = \min\{l \mid R_i^1 + \dots + R_i^{l-1} < j \text{ and } R_i^1 + \dots + R_i^l \geq j\}$ . Then take as item  $c_j$  any item in the database with value vector  $(x_1^{t_1(j)}, \dots, x_p^{t_p(j)})$ , and remove it from the database; the full supply assumption guarantees that it will always be possible to find such an item. Let  $A = \{c_1, \dots, c_k\}$ ; it is easy to check that  $A$  is an optimal committee for  $\|\cdot\|_1$  and for  $\|\cdot\|_{1,\max}$ .  $\square$

To illustrate the constructive proof, consider 2 attributes  $X_1$  with 3 values  $x_1^1, x_1^2, x_1^3$ , and  $X_2$  with 2 values  $x_2^1, x_2^2$ ;  $k = 4$ ; and  $R_1^1 = 2, R_1^2 = 0, R_1^3 = 2, R_2^1 = 3, R_2^2 = 1$ . Then  $t_1(1) = t_1(2) = 1, t_1(3) = t_1(4) = 3, t_2(1) = t_2(2) = t_2(3) = 1, t_2(4) = 2$ , which leads to choose  $c_1$  with value vector  $(x_1^1, x_2^1)$ ,  $c_2$  with vector  $(x_1^1, x_2^1)$ ,  $c_3$  with vector  $(x_1^3, x_2^1)$ , and  $c_4$  with vector  $(x_1^3, x_2^2)$ .

## 5 Properties of multi-attribute proportional representation

Several properties of apportionment methods have been studied, starting with Balinski and Young [1]. We omit their definition in the single-attribute case and directly give their generalizations to our more general model. Let  $A$  be any optimal committee for some criterion given  $\pi, C$  and  $k$ . We recall that  $R_i^j(A) = k r_i^j(A)$  denotes the number of elements of  $A$  with the attribute  $X_i$  equal to  $x_i^j$ .

- *Non-reversal*: for any attribute  $X_i$ , and attribute values  $x_i^j, x_i^{j'}$ , if  $\pi_i^j > \pi_i^{j'}$  then  $r_i^j(A) \geq r_i^{j'}(A)$ .
- *Exactness and respect of quota*: for all  $i$ , either  $R_i^j = \lfloor k\pi_i^j \rfloor$  or  $R_i^j = \lceil k\pi_i^j \rceil$ .
- *Population monotonicity* (with respect to  $X_i$ ): consider  $\pi$  and  $\rho$  such that (a)  $\pi_i^j > \rho_i^j$ , (b) for all  $j', j'' \neq j$ ,  $\frac{\pi_i^{j''}}{\pi_i^{j'}} = \frac{\rho_i^{j''}}{\rho_i^{j'}}$ , and (c) for all  $i' \neq i$  and all  $j$ ,  $\rho_i^j = \pi_i^j$ . Then there is an optimal committee  $B$  for  $\rho$  such that  $r_i^j(A) \geq r_i^j(B)$ .
- *House monotonicity*: let  $B$  be an optimal committee for  $\pi, C$  and  $k' > k$ . Then for all  $i, j$ ,  $r_i^j(B) \geq r_i^j(A)$ .<sup>4</sup>

In the single-attribute case, it is known for long that the Hamilton method satisfies all these properties except house monotonicity (this failure of house monotonicity is better known under the name *Alabama paradox*).

We start by noticing that if a property fails to be satisfied in the single-attribute case, *a fortiori* it is not satisfied in the multi-attribute case. As a consequence, house monotonicity is not satisfied, even under the FS assumption. We now consider the other properties.

<sup>4</sup>Some other properties, such as *consistency*, seem more difficult to generalize to the multi-attribute case. Also, properties that deal with strategy proofness issues, such as resistance to party merging or party splitting, are less relevant in our setting than for political elections and we omit them.

**Proposition 4** *Under the full supply assumption, non-reversal, exactness and respect of quota, and population monotonicity are all satisfied, for any of our loss functions. In the general case, non-reversal, exactness and respect of quota are not satisfied. If  $X_i$  is a binary variable, and for  $\|\cdot\|_1$ , population monotonicity with respect to  $X_i$  is satisfied; however it is not satisfied in the general case.*

*Proof.* Under FS, the result easily comes from Proposition 3 and the fact that the property holds in the single-attribute case.

In the general case, we give counterexamples. For exactness and respect of quota, we have two binary attributes, and two items  $a, b$  with value vectors  $(x_1^2, x_2^2)$  and  $(x_1^1, x_2^1)$ ,  $k = 1$ ,  $\pi$  defined as  $\pi_1^1 = 0, \pi_1^2 = 1, \pi_2^1 = 1, \pi_2^2 = 0$ . The optimal committee is either  $\{a\}$  or  $\{b\}$ , and does not respect quota even though all values  $k\pi_i^j$  are integers.

For non-reversal we have two binary attributes and six items:  $a, b, c$ , each with vector  $(x_1^1, x_2^1)$  and  $d, e, f$ , each with vector  $(x_1^2, x_2^2)$ . We have a target distribution  $\pi$  defined as follows:  $\pi_1^1 = 0.35, \pi_1^2 = 0.65, \pi_2^1 = 1, \pi_2^2 = 0$ . We set  $k = 3$ . The optimal committees for  $\|\cdot\|_1$  and  $\|\cdot\|_{1,\max}$  are  $\{a, b, c\}$  and all triples made up from two items out of  $\{a, b, c\}$  and one out of  $\{d, e, f\}$ . The optimal committees for  $\|\cdot\|_{\max}$  are all triples made up from two items out of  $\{a, b, c\}$  and one out of  $\{d, e, f\}$ . In all cases, for all optimal committees  $A$  we have  $r_1^1(A) > r_1^2(A)$  although  $\pi_1^1 < \pi_1^2$ .

Now, we prove that population monotonicity holds for binary domains and for  $\|\cdot\|_1$ . Consider a binary attribute  $X_i$ , with  $D_i = \{x_i^0, x_i^1\}$ .

Assume that  $\rho_i^0 > \pi_i^0$  (and  $\rho_i^0 > \pi_i^0$ ), and that for all  $i' \neq i$  we have  $\rho_{i'} = \pi_{i'}$ . Let  $A$  be an optimal committee for  $\pi$  and, for the sake of contradiction, assume that for all optimal committees  $B$  for  $\rho$  we have  $r_i^0(B) < r_i^0(A)$ . Let  $B$  be such a committee. The proof is a case by case study, with six cases to be considered: (C1)  $r_i^0(B) \leq \pi_i^0 < \rho_i^0 \leq r_i^0(A)$ ; (C2)  $\pi_i^0 \leq r_i^0(B) \leq \rho_i^0 \leq r_i^0(A)$ ; (C3)  $\pi_i^0 < \rho_i^0 \leq r_i^0(B) < r_i^0(A)$ ; (C4)  $r_i^0(B) \leq \pi_i^0 \leq r_i^0(A) \leq \rho_i^0$ ; (C5)  $\pi_i^0 \leq r_i^0(B) < r_i^0(A) \leq \rho_i^0$ ; and (C6)  $r_i^0(B) < r_i^0(A) \leq \pi_i^0 < \rho_i^0$ .

- Case 1:  $r_i^0(B) \leq \pi_i^0 < \rho_i^0 \leq r_i^0(A)$ . In this case we have  $r_i^1(B) \geq \pi_i^1 > \rho_i^1 \geq r_i^1(A)$  and the following holds:

$$\|r(B) - \pi\|_1 = \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \pi_{i'}^j| + (\pi_i^0 - r_i^0(B)) + (r_i^1(B) - \pi_i^1) \quad (1)$$

$$= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(B)) + (r_i^1(B) - \rho_i^1) + \pi_i^0 - \pi_i^1 - \rho_i^0 + \rho_i^1 \quad (2)$$

$$= \|r(B) - \rho\|_1 + 2(\pi_i^0 - \rho_i^0) \quad (3)$$

$$< \|r(A) - \rho\|_1 + 2(\pi_i^0 - \rho_i^0) \quad (4)$$

$$= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \rho_i^0) + (\rho_i^1 - r_i^1(A)) + 2(\pi_i^0 - \rho_i^0) \quad (5)$$

$$= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \pi_i^0) + (\pi_i^1 - r_i^1(A)) + \pi_i^0 - \pi_i^1 - \rho_i^0 + \rho_i^1 + 2(\pi_i^0 - \rho_i^0) \quad (6)$$

$$= \|r(A) - \pi\|_1 + 4(\pi_i^0 - \rho_i^0) \quad (7)$$

$$\leq \|r(A) - \pi\|_1 \quad (8)$$

(4) comes from the fact that  $A$  is not optimal for  $\rho$ . Since, there is one strong inequality in the sequence, we imply that  $A$  is not optimal for  $\pi$ , a contradiction.

- Case 2:  $\pi_i^0 \leq r_i^0(B) \leq \rho_i^0 \leq r_i^0(A)$ .



$$\begin{aligned}
\|r(B) - \pi\|_1 &= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \pi_{i'}^j| + (r_i^0(B) - \pi_i^0) + (\pi_i^1 - r_i^1(B)) \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(B)) + (r_i^1(B) - \rho_i^1) \\
&\quad + 2r_i^0(B) - \pi_i^0 - \rho_i^0 - 2r_i^1(B) + \pi_i^1 + \rho_i^1 \\
&= \|r(B) - \rho\|_1 + 4r_i^0(B) - 2\pi_i^0 - 2\rho_i^0 \\
&< \|r(A) - \rho\|_1 + 4r_i^0(B) - 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \rho_i^0) + (\rho_i^1 - r_i^1(A)) + 4r_i^0(B) - 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \pi_i^0) + (\pi_i^1 - r_i^1(A)) \\
&\quad + \pi_i^0 - \rho_i^0 - \pi_i^1 + \rho_i^1 + 4r_i^0(B) - 2\pi_i^0 - 2\rho_i^0 \\
&= \|r(A) - \pi\|_1 + 4r_i^0(B) - 4\rho_i^0 \\
&\leq \|r(A) - \pi\|_1
\end{aligned}$$

Again we obtain a contradiction.

- Case 3:  $\pi_i^0 < \rho_i^0 \leq r_i^0(B) < r_i^0(A)$ .

$$\begin{aligned}
\|r(B) - \pi\|_1 &= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \pi_{i'}^j| + (r_i^0(B) - \pi_i^0) + (\pi_i^1 - r_i^1(B)) \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \rho_{i'}^j| + (r_i^0(B) - \rho_i^0) + (\rho_i^1 - r_i^1(B)) \\
&\quad - \pi_i^0 + \rho_i^0 + \pi_i^1 - \rho_i^1 \\
&= \|r(B) - \rho\|_1 - 2\pi_i^0 + 2\rho_i^0 \\
&< \|r(A) - \rho\|_1 - 2\pi_i^0 + 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \rho_i^0) + (\rho_i^1 - r_i^1(A)) - 2\pi_i^0 + 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \pi_i^0) + (\pi_i^1 - r_i^1(A)) \\
&\quad + \pi_i^0 - \rho_i^0 - \pi_i^1 + \rho_i^1 - 2\pi_i^0 + 2\rho_i^0 \\
&= \|r(A) - \pi\|_1
\end{aligned}$$

- Case 4:  $r_i^0(B) \leq \pi_i^0 \leq r_i^0(A) \leq \rho_i^0$ .

$$\begin{aligned}
\|r(B) - \pi\|_1 &= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \pi_{i'}^j| + (\pi_i^0 - r_i^0(B)) + (r_i^1(B) - \pi_i^1) \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(B)) + (r_i^1(B) - \rho_i^1) \\
&\quad + \pi_i^0 - \rho_i^0 - \pi_i^1 + \rho_i^1 \\
&= \|r(B) - \rho\|_1 + 2\pi_i^0 - 2\rho_i^0 \\
&< \|r(A) - \rho\|_1 + 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(A)) + (r_i^1(A) - \rho_i^1) + 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \pi_i^0) + (\pi_i^1 - r_i^1(A)) \\
&\quad - 2r_i^0(A) + 2r_i^1(A) + \pi_i^0 + \rho_i^0 - \pi_i^1 - \rho_i^1 + 2\pi_i^0 - 2\rho_i^0 \\
&= \|r(A) - \pi\|_1 - 4r_i^0(A) + 4\pi_i^0 \\
&\leq \|r(A) - \pi\|_1
\end{aligned}$$

- Case 5:  $\pi_i^0 \leq r_i^0(B) < r_i^0(A) \leq \rho_i^0$ .

$$\begin{aligned}
\|r(B) - \pi\|_1 &= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \pi_{i'}^j| + (r_i^0(B) - \pi_i^0) + (\pi_i^1 - r_i^1(B)) \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(B)) + (r_i^1(B) - \rho_i^1) \\
&\quad + 2r_i^0(B) - 2r_i^1(B) - \pi_i^0 - \rho_i^0 + \pi_i^1 + \rho_i^1 \\
&= \|r(B) - \rho\|_1 + 4r_i^0(B) - 2\pi_i^0 - 2\rho_i^0 \\
&< \|r(A) - \rho\|_1 + 4r_i^0(B) - 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(A)) + (r_i^1(A) - \rho_i^1) + 4r_i^0(B) - 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (r_i^0(A) - \pi_i^0) + (\pi_i^1 - r_i^1(A)) \\
&\quad + 4r_i^0(B) - 2r_i^0(A) + 2r_i^1(A) + \pi_i^0 + \rho_i^0 - \pi_i^1 - \rho_i^1 - 2\pi_i^0 - 2\rho_i^0 \\
&= \|r(A) - \pi\|_1 + 4r_i^0(B) - 4r_i^0(A) \\
&\leq \|r(A) - \pi\|_1
\end{aligned}$$

- Case 6:  $r_i^0(B) < r_i^0(A) \leq \pi_i^0 < \rho_i^0$ .

$$\begin{aligned}
\|r(B) - \pi\|_1 &= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \pi_{i'}^j| + (\pi_i^0 - r_i^0(B)) + (r_i^1(B) - \pi_i^1) \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(B) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(B)) + (r_i^1(B) - \rho_i^1) \\
&\quad + \pi_i^0 - \rho_i^0 - \pi_i^1 + \rho_i^1 \\
&= \|r(B) - \rho\|_1 + 2\pi_i^0 - 2\rho_i^0 \\
&< \|r(A) - \rho\|_1 + 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (\rho_i^0 - r_i^0(A)) + (r_i^1(A) - \rho_i^1) + 2\pi_i^0 - 2\rho_i^0 \\
&= \sum_{i' \neq i} \sum_j |r_{i'}^j(A) - \rho_{i'}^j| + (\pi_i^0 - r_i^0(A)) + (r_i^1(A) - \pi_i^1) \\
&\quad - \pi_i^0 + \rho_i^0 + \pi_i^1 - \rho_i^1 + 2\pi_i^0 - 2\rho_i^0 \\
&= \|r(A) - \pi\|_1
\end{aligned}$$

Finally, we give an example showing that population monotonicity does not hold in the general case for  $\|\cdot\|_1$ . First, we describe the set of attributes. We have one distinguished attribute  $X_1$  with 5 possible values  $x_1^1, x_1^2, x_1^3, x_1^4$ , and  $x_1^5$  and 64 groups of binary attributes, indexed with the pairs of integers  $i, j \in \{1, 2, 3, 4\}$ . These groups of attributes are denoted as  $X_{(1,1)}, X_{(1,2)}, \dots, X_{(1,8)}, X_{(2,1)}, \dots, X_{(8,8)}$ . Each group contains some large number  $\lambda$  of indistinguishable attributes, each having the same set of possible values  $\{x_2^1, x_2^2\}$ . We have 16 alternatives  $A_1, A_2, \dots, A_8$ , and  $B_1, B_2, \dots, B_8$ , and our goal is to select a subset of  $k = 8$  of them.

We start with describing these alternatives on binary attributes: each alternative  $A_i$  has the value  $x_2^1$  on all attributes  $X_{(i,\cdot)}$  and the value  $x_2^2$  on all the remaining ones; each alternative  $B_i$  has the value  $x_2^1$  on all attributes  $X_{(\cdot,i)}$  and the value  $x_2^2$  on all the remaining ones. For the binary attributes we set the target probabilities to  $\pi_2^1 = 1/8$  and  $\pi_2^2 = 7/8$ . Due to this construction, we see that the only two subsets that perfectly agree with target distributions on each of binary attributes are  $A = \{A_1, A_2, \dots, A_8\}$  and  $B = \{B_1, B_2, \dots, B_8\}$ . Indeed, every subset  $S$  including  $A_i$  and  $B_j$ , would have  $r(S) \geq 1/4$  at least for one group of attributes  $X_{(i,j)}$ . Since  $\lambda$  is large, we infer that, independently what happens on the distinguished attribute  $X_1$ , the only possible winning committee is either  $A = \{A_1, A_2, \dots, A_8\}$  or  $B = \{B_1, B_2, \dots, B_8\}$ .

Next, let us describe what happens on the attribute  $X_1$ . The vector  $\langle r_i^j(A) \rangle$  is equal to  $\langle r_i^j(A) \rangle = (1/2, 0, 1/2, 0, 0)$ . For the committee  $B$ , we have  $\langle r_i^j(B) \rangle = (1/4, 1/4, 1/4, 1/8, 1/8)$ , and the vector of target distributions for  $X_1$  is equal  $\pi_1 = (0, 0, 3/8 + \epsilon, 5/8 - \epsilon, 0)$ . We can see that  $\|r(A) - \pi\|_1 = 1/2 + 1/8 - \epsilon + 5/8 - \epsilon = 1.25 - 2\epsilon$ . Since,  $\|r(B) - \pi\|_1 = 1/4 + 1/4 + 1/8 + \epsilon + 4/8 - \epsilon + 1/8 = 1.25$ , we get that  $A$  is a winning committee. However, if we modify the target fractions so that  $\rho_1 = (1/4, 0, 9/32 + \epsilon_1, 15/32 - \epsilon_2, 0)$ , we will get  $\|r(A) - \rho\|_1 = 1/4 + 7/32 - \epsilon_1 + 15/32 - \epsilon_2 = 30/32 - \epsilon_1 - \epsilon_2$  and  $\|r(B) - \rho\|_1 = 1/4 + 1/32 + \epsilon_1 + 11/32 - \epsilon_2 + 1/8 = 24/32 + \epsilon_1 - \epsilon_2$ , thus,  $B$  is winning according to  $\rho$ . However,  $B$  has lower representation of  $x_1^1$  than  $A$ , and  $\rho$  was obtained from  $\pi$ , by increasing the fraction of  $\pi_1^1$ . This completes the proof.  $\square$

Other properties, specific to multi-attribute proportional representation, could also be considered, for instance by adapting properties studied by Elkind et al. [10]. One such property is *candidate monotonicity* (if we add more candidates to the database, the new committee must be at least as good as the old one). We leave this for further research.

## 6 Computing Optimal Committees

In this section we now investigate the computation complexity of optimal committees. We start with observing that the problem of deciding whether there is a perfect committee for a given instance is **NP**-complete.

**Proposition 5** *Given set of attributes  $X$ , a set of candidates  $C$ , a vector of target distributions  $\pi$ , an integer  $k$ , deciding whether there is a perfect committee is **NP**-complete.*

*Proof.* Membership is straightforward. Hardness follows by reduction from the **NP**-complete problem EXACT COVER WITH 3-SETS, or X3C [12]. Let  $I = \langle X, \mathcal{S} \rangle$  with  $X = \{x_1, \dots, x_{3k}\}$  and  $\mathcal{S} = \{S_1, \dots, S_n\}$

with  $|S_i| = 3$  for each  $i$ .  $I$  is a positive instance of  $\text{x3C}$  iff there is a collection  $\mathcal{S}' \subseteq \mathcal{S}$  with  $|\mathcal{S}'| = k$  and  $\cup\{S | S \in \mathcal{S}'\} = X$ . Define the following instance of  $\text{PERFECT COMMITTEE}$ : let  $X_1, \dots, X_{3k}$  be  $3k$  binary attributes, and let  $C$  consist of  $m$  candidates  $c_1, \dots, c_m$  with  $X_i(c_j) = 1$  if  $x_i \in S_j$  and  $X_i(c_j) = 0$  if  $x_i \notin S_j$ . Finally, for each  $i$ ,  $\pi_i(0) = \frac{k-1}{k}$  and  $\pi_i(1) = \frac{1}{k}$ . We want a committee of size  $k$ .  $A = \{c_{i_1}, \dots, c_{i_k}\}$  is perfect for  $\pi$  if for each  $X_i$ , there is exactly one  $j \in \{1, \dots, k\}$  such that  $X_i(c_{i_j}) = 1$ , which is equivalent to saying that for each  $x_i$ , there is exactly one  $S_j \in \{S_{i_1}, \dots, S_{i_k}\}$  such that  $x_i \in S_j$ . Thus, there is a perfect committee for  $\pi$  and  $C$  if and only if  $I$  is a positive instance.  $\square$

This simple result implies that the decision problem associated with finding an optimal committee (is there a committee whose loss is less than  $\theta$ ?) is  $\text{NP-hard}$  for *all* loss functions. However, if the number of attributes  $p$  is fixed, the problem is solvable in polynomial time.

**Proposition 6** *Let  $p$  be a constant integer. Given set of  $p$  attributes  $X$ , a set of candidates  $C$ , a vector of target distributions  $\pi$ , an integer  $k$ , deciding whether there is a perfect committee is solvable in polynomial time.*

*Proof.* Let  $q = \max_i q_i$ . Each candidate can be viewed as a vector of values indexed with the attributes; there are  $q^p$  such possible vectors. Since the size of the input is at least  $q$ , the number of distinct candidates is bounded by the polynomial function of the size of the input. The rest of the proof is the same as the proof of Theorem 4.  $\square$

## 6.1 Approximating optimal committees

A natural approach to alleviate the  $\text{NP-hardness}$  of the problem is to analyze whether it can be well approximated. Before proceeding to presentation of our approximation algorithms, the core technical contribution of this paper, we define the notion of approximability used in our analysis.

**Definition 6** *An algorithm  $\mathcal{A}$  is an  $\alpha$ -additive-approximation algorithm for  $\text{OPTIMALREPRESENTATION}$  if for each instance  $I$  of  $\text{OPTIMALREPRESENTATION}$  it holds that  $|f(\pi, r(A)) - f(\pi, r(A^*))| \leq \alpha$ , where  $A$  is the committee returned by  $\mathcal{A}$  for  $I$ , and  $A^*$  an optimal committee.*

It is easy to observe that for binary domains it holds that  $\|\pi, r(A)\|_1 = 2\|\pi, r(A)\|_{1, \max}$ . This implies that for binary domains, an  $\alpha$ -additive-approximation algorithm for  $\|\cdot\|_1$  is an  $\frac{\alpha}{2}$ -additive-approximation algorithm for  $\|\cdot\|_{1, \max}$ .

In this paper we mostly present computational results for binary domains. However, this assumption is not as restrictive as it may seem—every instance of the  $\text{OPTIMALREPRESENTATION}$  problem can be transformed to a new instance with binary domains in the following way:

- $X_{\text{new}} = \{X_{ij} \mid i = 1, \dots, p, j = 1, \dots, |D_i|\}$ .
- $C_{\text{new}} = \{c'_l \mid l = 1, \dots, m\}$ .
- $\pi_{\text{new}} = (\pi_{i,j} \mid 1 \leq i \leq p, 1 \leq j \leq |D_i|)$ , where for all  $i = 1, \dots, m, j = 1, \dots, p$  and  $j = 1, \dots, |D_i|$ ,  $\pi_{i,j}^0 = \pi_i^j$  and  $\pi_{i,j}^1 = 1 - \pi_i^j$ .

The following lemma shows how to obtain approximation guarantees for arbitrary domains having guarantees for the problem transformed to binary domains.

**Lemma 1** *For a given committee  $A$  and target distribution  $\pi$ , let  $A_{\text{new}}$  and  $\pi_{\text{new}}$  denote the committee and target distributions obtained as above. The following holds:*

$$I. \quad \|\pi_{\text{new}}, r(A_{\text{new}})\|_1 = 2\|\pi, r(A)\|_1.$$

2.  $1 \leq \frac{\|\pi_{\text{new}}, r(A_{\text{new}})\|_{1, \max}}{\|\pi, r(A)\|_{1, \max}} \leq \max_i |D_i|$ .
3.  $\max(\pi_{\text{new}}, r(A_{\text{new}})) = \max(\pi, r(A))$ .

*Proof.* We prove the first equality—the proof for the other two is similar.

$$\begin{aligned}
\|\pi, r(A)\|_1 &= \sum_{i,j} |r_i^j(A) - \pi_i^j| = \sum_{i,j} \left| \frac{|\{c \in A : X_i(c) = x_i^j\}|}{k} - \pi_i^j \right| \\
&= \sum_{i,j} \left| \frac{|\{c \in A_{\text{new}} : X_{i,j}(c) = 1\}|}{k} - \pi_{i,j}^1 \right| \\
&= \frac{1}{2} \sum_{i,j} \left( \left| \frac{|\{c \in A_{\text{new}} : X_{i,j}(c) = 1\}|}{k} - \pi_{i,j}^1 \right| + \left| \frac{|\{c \in A_{\text{new}} : X_{i,j}(c) = 0\}|}{k} - \pi_{i,j}^0 \right| \right) \\
&= \frac{1}{2} \sum_{i,j} \sum_{\ell \in \{0,1\}} |r_{i,j}^\ell(A) - \pi_{i,j}^\ell| = \frac{1}{2} \|\pi_{\text{new}}, r(A_{\text{new}})\|_1.
\end{aligned}$$

□

Lemma 1 has interesting implications—first shows that the transformed instance has the same perfect committees as the original instance; then it shows how to obtain additive approximation guarantees for arbitrary domains having guarantees for the problem restricted to binary domains, for different loss functions.

## 6.2 Approximation algorithms

In this section we show an approximation algorithm for the OPTIMALREPRESENTATION problem. The algorithm is given in Figure 1 and is parameterized by an integer value  $\ell$ . It starts with a random collection of  $k$  samples and, in each step, it looks whether it is possible to replace some  $\ell$  items from the current solution with some other  $\ell$  items to obtain a better solution. The algorithm continues until it cannot find any pair of sets of  $\ell$  items that improves the current solution. As we show now, the approximation guarantees depend on the value of the parameter  $\ell$ .

### Parameters:

$\pi = (\pi_1, \dots, \pi_p)$ —input target distributions.

$\ell$ —the parameter of the algorithm.

$A \leftarrow k$  random items from  $C$ ;

**while** there exist  $C_\ell \subset C$  and  $A_\ell \subset A$  such that  $|C_\ell| \leq \ell$ ,  $|A_\ell| \leq \ell$ , and  $f(\pi, r(A)) > f(\pi, r((A \setminus A_\ell) \cup C_\ell))$

**do**

$A \leftarrow (A \setminus A_\ell) \cup C_\ell$ ;

**return**  $A$ ;

Figure 1: Local search approximation algorithm.

**Theorem 1** *For binary domains natural distributions, and for the  $\|\cdot\|_1$  loss function, the local search algorithm defined on Figure 1 with  $\ell = 1$  is a  $|X|$ -additive-approximation algorithm for OPTIMALREPRESENTATION.*

*Proof.* Let  $A^*$  denote an optimal solution for a given instance  $I$  of the problem of finding a perfect committee. Let  $A \in S_k(C)$  denote the set returned by the local search algorithm from Figure 1. From the condition in the “while” loop, we know that there exist no  $c \in C$  and  $a \in A$  such that  $\|\pi, r(A)\|_1 > \|\pi, r((A \setminus \{a\}) \cup \{c\})\|_1$ .

Now, let  $X_{\text{ex}} \subseteq X$  denote the set of all attributes for which  $A$  achieves exact match with  $\pi$ , that is, such that for each  $X_i \in X_{\text{ex}}$ , we have that  $r_i^1(A) = \pi_i^1$  and  $r_i^2(A) = \pi_i^2$ .

Let us consider the procedure consisting in taking the items from  $A \setminus A^*$  and, one by one, replace them with arbitrary items from  $A^* \setminus A$ . This procedure, in  $|A \setminus A^*|$  steps, transforms  $A$  into an optimal solution  $A^*$ . We now estimate the total gain  $g$  induced by this procedure. For each item  $a \in A \setminus A^*$ , by  $a' \in A^* \setminus A$  we denote the item which was taken to replace  $a$  in the procedure. For each attribute  $X_i \in X$  we define the gain  $g_i(a, a')$  of replacing  $a$  by  $a'$  as:

$$g_i(a, a') = \sum_{j \in \{1, 2\}} \left( |r_i^j(A) - \pi_i^j| - |r_i^j(A \setminus \{a\} \cup \{a'\}) - \pi_i^j| \right).$$

We now extend this definition to sets of  $k$  candidates:

$$g_i(B, B') = \sum_{j \in \{1, 2\}} \left( |r_i^j(A) - \pi_i^j| - |r_i^j((A \setminus B) \cup B') - \pi_i^j| \right).$$

If  $X_i \in X_{\text{ex}}$ , then  $r_i(A) = \pi_i$ , and so the replacement cannot improve the quality of the solution relatively to  $X_i$ , hence

$$\sum_{i \in X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) \leq 0. \quad (1)$$

Note that  $g_i(a, a') \in \{-\frac{2}{k}, 0, \frac{2}{k}\}$ . Moreover, for each attribute  $X_i \notin X_{\text{ex}}$  there are two possible cases:

1.  $r_i^j(A) > \pi_i^j$  and each exchange of candidate that results in a negative gain increases  $r_i^j(A)$ .
2.  $r_i^j(A) < \pi_i^j$  and each exchange that results in a negative gain decreases  $r_i^j(A)$ .

Intuitively, 1. and 2. mean that for attributes outside of  $X_{\text{ex}}$ , the negative gains cumulate. Formally, for each  $X \notin X_{\text{ex}}$ :

$$g_i(A \setminus A^*, A^* \setminus A) \leq \sum_{a \in A \setminus A^*} g_i(a, a'). \quad (2)$$

From the condition in the “while” loop, we have that for each  $a \in A \setminus A^*$ :  $\sum_i g_i(a, a') \leq 0$ , and so:

$$\sum_i \sum_{a \in A \setminus A^*} g_i(a, a') \leq 0. \quad (3)$$

We now give the following sequence of inequalities:

$$\begin{aligned} g &= \sum_i g_i(A \setminus A^*, A^* \setminus A) \\ &= \sum_{i \in X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) + \sum_{i \notin X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) \\ &\leq \sum_{i \notin X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) \leq \sum_{i \notin X_{\text{ex}}} \sum_{a \in A \setminus A^*} g_i(a, a') \\ &\leq - \sum_{i \in X_{\text{ex}}} \sum_{a \in A \setminus A^*} g_i(a, a') \\ &\leq |X_{\text{ex}}| \cdot k \cdot \frac{2}{k} = 2|X_{\text{ex}}|. \end{aligned} \quad (4)$$

Finally, for each attribute  $X_i \in X_{\text{ex}}$  the loss relative to  $X_i$ , i.e.,  $|r_i^0 - \pi^0| + |r_i^1 - \pi^1|$ , is at most 2. Thus, we get  $g \leq 2(|X| - |X_{\text{ex}}|)$ , which leads to  $g \leq |X|$ .  $\square$

Is the bound  $|X|$  from Theorem 1 a good result? One way to interpret this result is to observe that a solution that for half of the attributes gives exact match, and for other half is arbitrarily bad, is an  $|X|$ -approximate solution. We do not know whether the bound  $|X|$  is reached, but we now show that a lower bound on the error made by the algorithm with  $\ell = 1$  is  $\frac{2}{3}|X|$ .

**Example 2** Consider  $3p$  binary attributes  $X_1, \dots, X_{3p}$ ,  $4\ell$  candidates  $C = \{a_1, \dots, a_{2\ell}, b_1, \dots, b_{2\ell}\}$ , and let  $k = 2\ell$ . For each  $i \leq p$ , we have: for  $j \leq \ell$ ,  $X_i(a_j) = 1$  and  $X_i(b_j) = 1$ ; for  $j > \ell$ ,  $X_i(a_j) = 0$  and  $X_i(b_j) = 0$ . For each  $i$  such that  $p < i \leq 2p$  we have: for  $j \leq \ell$ ,  $X_i(a_j) = 1$  and  $X_i(b_j) = 0$ ; for  $j > \ell$ ,  $X_i(a_j) = 0$  and  $X_i(b_j) = 1$ . For  $i > 2p$  we have: for each  $j$ ,  $X_i(a_j) = 1$  and  $X_i(b_j) = 0$ . Finally, for  $i \leq 2p$  let  $\pi_i^0 = \pi_i^1 = \frac{1}{2}$ , and for  $i > 2p$  let  $\pi_i^0 = 1 - \pi_i^1 = 1$ . It can be easily checked that  $B = \{b_1, \dots, b_{2\ell}\}$  is a perfect committee. Now,  $A = \{a_1, \dots, a_{2\ell}\}$  is locally optimal. To check this, we consider two cases: in the first case, where  $(r \leq \ell \text{ and } q \leq \ell)$  or  $(r > \ell \text{ and } q > \ell)$ , replacing  $a_r$  with  $b_q$  does not change the distance to the target distribution on each of the first  $p$  attributes, increases the distance on each of the next  $p$  attributes and decreases the distance on each of the last  $p$  attributes. For the second case, where  $(r \leq \ell \text{ and } q > \ell)$  or  $(r > \ell; q \leq \ell)$ , the line of reasoning is similar. Finally,  $\|\pi, r(A)\|_1 = 2p = \frac{2}{3}|X|$ .

A better approximation bound can be obtained with  $\ell = 2$ :

**Lemma 2** Consider  $n$  buckets  $X_1, \dots, X_n$ , such that in the  $i$ -th bucket  $X_i$  there are  $x_i$  white balls and  $y_i$  black balls. Let  $A$  denote the number of pairs of balls such that both balls in the pair belong to the same bucket and are of different color. Let us consider the procedure in which one iteratively selects a bucket and takes out two balls with different colors from the selected bucket. The procedure ends after  $B$  steps, when no further steps are possible (in each bucket, either there are no balls anymore, or all balls have the same color). It holds that  $A \geq \frac{B^2}{n}$ .

*Proof.* Without loss of generality let us assume that for each  $i$ :  $x_i \leq y_i$ . Thus,  $B = \sum_i x_i$  and  $A = \sum_i x_i y_i \leq \sum_i x_i^2$ . The inequality  $\sum_i x_i^2 \geq \frac{(\sum_i x_i)^2}{n}$  follows from Jensen's inequality applied to the quadratic function.  $\square$

**Lemma 3** Let  $x_i, y_i, A_i$ ,  $1 \leq i \leq n$ , be real values satisfying the following constraints:

1.  $x_i \geq \frac{A_i}{2n-2(i-1)}$ , for each  $1 \leq i \leq n$ ,
2.  $A_i \geq A_{i-1} - 2x_{i-1}$ , for each  $2 \leq i \leq n$ ,
3.  $y_i \geq \frac{x_i}{2n-2(i-1)-1}$ , for each  $1 \leq i \leq n$ .

Then:

$$\sum_{i=1}^n y_i \geq \frac{|A_1| \ln n}{4n}.$$

*Proof.* We can view the set of above inequalities 1, 2, 3 as a linear program with  $(3n - 1)$  variables (all  $x_i$  and  $y_i$  for  $1 \leq i \leq n$  and  $A_i$  for  $2 \leq i \leq n$ ; we treat  $A_1$  as a constant) and  $(3n - 1)$  constraints. Thus, we know that  $\sum_i y_i$  achieves the minimum when each from the above constraints is satisfied with equality.

We show by induction that the values  $x_i = \frac{A_1}{2n}$  and  $A_i = \frac{2n-2(i-1)}{2n} A_1$  constitute the solution to the set of equalities that is derived by taking constraints 1, and 2, and treating them as equalities. We can show that

by induction: It is easy to see that the base step, for  $i = 1$ , holds:

$$x_1 = \frac{A_1}{2n - 2(i - 1)} = \frac{|A_1|}{2n},$$

$$A_1 \geq \frac{2n - 2(1 - 1)}{2n} A_1.$$

Let us assume that from the equalities 1 and 2 taken for  $i < j$ , it follows that  $x_i = \frac{A_1}{2n}$  and  $A_i = \frac{2n - 2(i - 1)}{2n} A_1$ , for  $i < j$ . We will show that from equalities 1 and 2 for  $i = j$  it follows that  $x_j = \frac{A_1}{2n}$  and  $A_j = \frac{2n - 2(j - 1)}{2n} A_1$ :

$$x_j = \frac{A_j}{2n - 2(j - 1)} = \frac{1}{2n - 2(j - 1)} \cdot \frac{2n - 2(j - 1)}{2n} A_1 = \frac{|A_1|}{2n},$$

$$A_j = A_{j-1} - 2x_{j-1} = \frac{2n - 2((j - 1) - 1)}{2n} A_1 - 2 \frac{|A_1|}{2n} = \frac{2n - 2(j - 1)}{2n} A_1.$$

From constraint 3, treated as equality, we get:

$$y_i = \frac{x_i}{2n - 2(i - 1) - 1} = \frac{|A_1|}{2n(2n - 2(i - 1) - 1)}.$$

Thus, we infer that  $\sum_{i=1}^n y_i$  is minimized when  $y_i = \frac{|A_1|}{2n(2n - 2(i - 1) - 1)}$ . We recall that  $H_n$  denotes the  $n$ -th harmonic number ( $H_n = \sum_{i=1}^n \frac{1}{i}$ ), and that  $\ln(n + 1) < H_n \leq 1 + \ln(n)$ . As a result we get:

$$\sum_{i=1}^n y_i \geq \frac{A_1}{2n} \sum_{i=1}^n \frac{1}{(2n - 2(i - 1) - 1)} \geq \frac{A_1}{2n} \sum_{i=1}^n \frac{1}{2n - 2(i - 1)} \quad (5)$$

$$= \frac{A_1}{4n} \sum_{i=1}^n \frac{1}{(n - i + 1)} = \frac{A_1}{4n} H_n \geq A_1 \frac{\ln n}{4n}. \quad (6)$$

□

**Theorem 2** For binary domains ( $|D_i| = 2$ , for each  $1 \leq i \leq p$ ), natural distributions, and for  $\|\cdot\|_1$  loss function, the local search algorithm from Figure 1 with  $\ell = 2$  is a  $\frac{\ln(k/2)}{2\ln(k/2)-1} \left(|X| + \frac{6|X|}{k}\right)$ -additive-approximation algorithm for OPTIMALREPRESENTATION.

*Proof.* In this proof we use similar idea to the proof of Theorem 1, but the proof is technically more involved. As before, by  $A^*$  and  $A$  we denote the optimal solution and the solution returned by the local search algorithm, respectively. Similarly to the previous proof, by  $X_{\text{ex}} \subset X$  we denote the set of all attributes for which  $A$  achieves exact match with  $\pi$ , i.e.,

$$X_{\text{ex}} = \{X_i \in X : r_i^1(A) = \pi_i^1\}.$$

We also define the set  $X_{\text{aex}} \subset X$  of all attributes for which  $A$  achieves almost exact match with  $\pi$ , i.e.,

$$X_{\text{aex}} = \left\{X_i \in X : |r_i^1(A) - \pi_i^1| \leq \frac{1}{k}\right\}.$$

Let  $q_f = \frac{|A \setminus A^*|}{2}$  and  $q = \lfloor q_f \rfloor$ . Let us rename the items from  $A \setminus A^*$  so that  $A \setminus A^* = \{a_1, a_2, \dots, a_{2q_f}\}$ , and the items from  $A^* \setminus A$ , so that  $A^* \setminus A = \{a'_1, a'_2, \dots, a'_{2q_f}\}$ . Hereinafter, we follow a convention in which the elements from  $A^* \setminus A$  are marked with primes. Renaming of the items that we described above, allows us to define the following sequence of pairs  $(a_1, a'_1), \dots, (a_{2q_f}, a'_{2q_f})$  in which each element from  $A \setminus A^*$  is paired with (assigned to) exactly one element from  $A^* \setminus A$ .

For each pair  $(a_j, a'_j)$  and for each attribute  $X_i$  we consider what happens if we replace  $a_i$  in  $A \setminus A^*$  with  $a'_i$ . One of three scenarios can happen, after such replacement:

1. The value  $r_i^0(A)$  can increase by  $\frac{1}{k}$  (in such case  $r_i^1(A)$  decreases by  $\frac{1}{k}$ ), which we denote by  $X_i(a_j \leftrightarrow a'_j) = 1$ ,
2. The value  $r_i^0(A)$  can decrease by  $\frac{1}{k}$  (in such case  $r_i^1(A)$  increases by  $\frac{1}{k}$ ), which we denote by  $X_i(a_j \leftrightarrow a'_j) = -1$ , or
3. The value  $r_i^0(A)$  can remain unchanged (in such case  $r_i^1(A)$  also remains unchanged), which we denote by  $X_i(a_j \leftrightarrow a'_j) = 0$ .

We follow a procedure which, in  $q$  consecutive steps, replaces pairs of items from  $A \setminus A^*$ , with the pairs of items from  $A^* \setminus A$ . A pair  $(a_i, a_j)$  is always replaced with  $(a'_i, a'_j)$ . In other words, when looking for a pair from  $A^* \setminus A$  to replace  $(a_i, a_j)$  we follow the assignment rule induced by renaming, as described above. The way in which we create pairs within  $A \setminus A^*$  for replacement (the way how  $(a_i, a_j)$  is selected in each of  $q$  consecutive steps) will be described later. After this whole procedure  $A$  can differ from  $A^*$  with at most one element, hence, having distance to the optimal distribution at most equal to  $|X| \frac{2}{k}$ . Let us define the sequence of sets  $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_q$  in the following way: we define  $\bar{A}_1 = A \setminus A^*$ , and we define  $\bar{A}_{j+1}$  as  $\bar{A}_j$  after removing the pair from  $A \setminus A^*$  that was used in replacement in the  $j$ -th step of our procedure.

As before, for each  $B \subseteq A \setminus A^*$  and  $B' \subseteq A^* \setminus A$ , and for each attribute  $X_i \in X$  we define the gain  $g_i(B, B')$ :

$$g_i(B, B') = \sum_{j \in \{1, 2\}} \left( |r_i^j(A) - \pi_i^j| - |r_i^j((A \setminus B) \cup B') - \pi_i^j| \right).$$

Similarly as in the proof of Theorem 1, we observe that for  $X_i \notin X_{\text{aex}}$  the negative gains cumulate: i.e., that for each sequences of disjoint sets  $B_1, B_2, \dots, B_s$  and  $B'_1, B'_2, \dots, B'_s$  such that for every  $1 \leq j \leq s$ ,  $B_j \subseteq A \setminus A^*$ ,  $B'_j \subseteq A^* \setminus A$ , and  $|B_j| = |B'_j| \leq 2$  we have that:

$$g_i\left(\bigcup_j B_j, \bigcup_j B'_j\right) \leq \sum_j g_i(B_j, B'_j). \quad (7)$$

Why is this the case? If  $X_i \notin X_{\text{aex}}$ , then the distance between  $A$  and the target distribution on attribute  $X_i$  is at least equal to  $2 \cdot \frac{2}{k}$ . In other words:  $|r_i^0(A) - \pi_i^0| \geq \frac{2}{k}$  and  $|r_i^1(A) - \pi_i^1| \geq \frac{2}{k}$ . Without loss of generality let us assume that  $r_i^0(A) - \pi_i^0 \geq \frac{2}{k}$ . Since each set  $B_j$  and each set  $B'_j$  has at most two elements, replacing  $B_j$  with  $B'_j$  can change the distance between  $A$  and the target distribution, for each attribute, by at most  $\frac{2}{k}$ . Consequently, if  $g_i(B_j, B'_j)$  is negative, then it means that replacing  $B_j$  with  $B'_j$  makes the difference  $r_i^0(A) - \pi_i^0$  even greater. Thus, each such replacement with the negative gain  $g$  causes  $A$  to move further from the target distribution by the value  $g$ . Naturally, each replacement with the positive gain  $g$  causes  $A$  to move closer to the target distribution by at most  $g$ . Consequently, after the sequence of replacement  $\bigcup_j B_j \leftrightarrow \bigcup_j B'_j$  the distance on the attribute  $X_i$  cannot improve by more than  $\sum_j g_i(B_j, B'_j)$ .

In contrast to the proof of Theorem 1, we note that here we require that  $X_i \notin X_{\text{aex}}$  instead of  $X_i \notin X_{\text{ex}}$ —the above observation is not valid if  $X_i \in X_{\text{aex}}$  even if  $X_i \notin X_{\text{ex}}$ .<sup>5</sup>

Next, for each  $\bar{A}_j$ , and each attribute  $X_i \in X_{\text{ex}}$ , we define a set  $W_j$  of annihilating pairs as:

$$W_j(X_i) = \{((a_x, X_i), (a_y, X_i)) : a_x \in \bar{A}_j; a_y \in \bar{A}_j; x < y; X_i(a_x \leftrightarrow a'_x) = -X_i(a_y \leftrightarrow a'_y)\}.$$

<sup>5</sup> Consider an example in which  $\pi_i^1 = \frac{1}{k}$  and  $r_i^1(A) = \frac{2}{k}$ . Let us consider sets  $B = \{b_1, b_2\}$ ,  $B' = \{b'_1, b'_2\}$ ,  $C = \{c_1, c_2\}$ ,  $C' = \{c'_1, c'_2\}$  such that:  $X_i(c_1) = X_i(c_2) = X_i(b'_1) = X_i(b'_2) = d_i^1$ , and  $X_i(c'_1) = X_i(c'_2) = X_i(b_1) = X_i(b_2) = d_i^2$ . Thus, we have that:

- Replacing  $B$  with  $B'$  results with  $r_i^1(A) = \frac{4}{k}$ .
- Replacing  $C$  with  $C'$  results with  $r_i^1(A) = 0$ .
- Replacing  $B \cup C$  with  $B' \cup C'$  results with  $r_i^1(A) = \frac{2}{k}$ .

We can repeat this reasoning for  $r_i^2(A)$ , thus having,  $g_i(B, B') = -\frac{4}{k}$ ,  $g_i(C, C') = 0$  and  $g_i(B \cup C, B' \cup C') = 0$ .



	$X_i = X_1$	$X_i = X_2$	$X_i = X_3$	$X_i = X_4$	$X_i = X_5$	$X_i = X_6$	$X_i = X_7$
$X_i(a_1 \leftrightarrow a'_1)$	1	1	1	1	0	0	-1
$X_i(a_2 \leftrightarrow a'_2)$	-1	-1	1	0	0	1	0
$X_i(a_3 \leftrightarrow a'_3)$	0	-1	-1	0	1	0	1
$X_i(a_4 \leftrightarrow a'_4)$	-1	1	-1	-1	1	0	-1

Table 1: An example illustrating the concept of annihilating pairs. In this example we have  $X_{\text{ex}} = \{X_1, X_2, X_3, X_4, X_5, X_6, X_7\}$  and  $\bar{A}_1 = \{a_1, a_2, a_3, a_4\}$ . We recall that  $X_i(a_i \leftrightarrow a'_i) = 1$  if replacing  $a_i$  with  $a'_i$  moves  $A$  further from the target distribution in one direction and  $X_i(a_i \leftrightarrow a'_i) = -1$  if replacing  $a_i$  with  $a'_i$  moves  $A$  further from the target distribution in the other direction. Here, we have  $W_1(X_1) = \{((a_1, X_1), (a_2, X_1)), ((a_1, X_1), (a_4, X_1))\}$ ,  $W_1(X_2) = \{((a_1, X_2), (a_2, X_2)), ((a_1, X_2), (a_3, X_2))\}$ ,  $W_1(X_3) = \{((a_1, X_3), (a_3, X_3)), ((a_1, X_3), (a_4, X_3)), ((a_2, X_3), (a_3, X_3)), ((a_2, X_3), (a_4, X_3))\}$ , etc. Further,  $W_1 = W_1(X_1) \cup W_1(X_2) \cup W_1(X_3) \cup W_1(X_4) \cup W_1(X_5) \cup W_1(X_6) \cup W_1(X_7)$ . There are many choices for the set  $P$ , but it must hold that  $|P| = 6$ ; we give the following example:  $P = \{((a_1, X_1), (a_2, X_1)), ((a_1, X_2), (a_2, X_2)), ((a_1, X_3), (a_3, X_3)), ((a_2, X_3), (a_4, X_3)), ((a_1, X_4), (a_4, X_4)), ((a_1, X_7), (a_3, X_7))\}$ .

Intuitively, if  $((a_x, X_i), (a_y, X_i)) \in W_j$ , then both replacing  $a_x$  with  $a'_x$  and replacing  $a_y$  with  $a'_y$  move the original set  $A$  (i.e., the set before any of the replacements) further from the target distribution for the attribute  $X_i$ , but replacing  $\{a_x, a_y\}$  with  $\{a'_x, a'_y\}$  does not change the distance of  $A$  from the target distribution for the attribute  $X_i$ .

For each  $j$ , we set  $W_j = \cup_{i \in X_{\text{ex}}} W_j(X_i)$ . Let us denote by  $P$  the number of annihilated pairs of candidates considered in the process of replacing items from  $A \setminus A^*$  with items from  $A^* \setminus A$ . Formally,  $P$  is the size of the maximal subset  $W \subseteq W_1$  composed of disjoint annihilating pairs, i.e., for each  $i \leq p$ , for each  $a_x$ , and for each  $a_y$ , if  $((a_x, X_i), (a_y, X_i)) \in P$  then there exists no  $b \neq a_y$  such that  $((a_x, X_i), (b, X_i)) \in P$  or  $((b, X_i), (a_x, X_i)) \in P$ . From Lemma 2, after defining each bucket  $X_i$  as containing  $x_i$  white balls and  $y_i$  black balls, where  $x_i$  (respectively,  $y_i$ ) is the number of candidates  $a_j \in \bar{A}_1$  with the value  $X_i(a_j \leftrightarrow a'_j)$  equal to 1 (respectively, -1), it follows that  $W_1 \geq \frac{P^2}{|X_{\text{ex}}|}$ . The concept of annihilating pairs is explained on example in Table 1.

We are now ready to describe the way in which we select pairs from  $A \setminus A^*$  in our procedure. In each step  $j$ , the pair  $(a_{j,1}, a_{j,2})$  from  $A \setminus A^*$  is selected in the following way. For each item  $a$  let  $s_{j,1}(a)$  be the number of pairs  $p$  in  $W_j$  such that  $p = ((a, \cdot), (\cdot, \cdot))$  or  $p = ((\cdot, \cdot), (a, \cdot))$ , let  $a_{j,1}$  be such that  $s_{j,1}(a_{j,1}) = \max_{a \in \bar{A}_j} s_{j,1}(a)$ , and let  $s_{j,1} = s_{j,1}(a_{j,1})$ . Next, for each item  $b$  let  $s_{j,2}(b)$  be the number of pairs  $p$  in  $W_j$  such that  $p = ((a_{j,1}, \cdot), (b, \cdot))$  or  $p = ((b, \cdot), (a_{j,1}, \cdot))$ , let  $a_{j,2}$  be such that  $s_{j,2}(a_{j,2}) = \max_{b \in \bar{A}_j} s_{j,2}(b)$ , and let  $s_{j,2} = s_{j,2}(a_{j,2})$ .

Let us consider the procedure described above on the example from Table 1. The item  $a_1$  belongs to 8 pairs in  $W_1$  ( $a_1$  belongs to 2 pairs for attribute  $X_1, X_2$ , and  $X_3$ , and to one pair for attributes  $X_4$  and  $X_7$ ), thus:  $s_{1,1}(a_1) = 8$ . Moreover,  $s_{1,1}(a_2) = 5$ ,  $s_{1,1}(a_3) = 6$ , and  $s_{1,1}(a_4) = 7$ . Consequently,  $a_1$  will be the item that will be replaced with  $a'_1$  in the first step:  $a_{j,1} = a_1$  and  $s_{j,1} = 8$ . Further,  $s_{1,2}(a_2) = 2$  (there are two annihilating pairs including  $a_1$  and  $a_2$ , i.e.,  $((a_1, X_1), (a_2, X_1))$  and  $((a_1, X_2), (a_2, X_2))$ ); similarly:  $s_{1,2}(a_3) = 3$ , and  $s_{1,2}(a_4) = 3$ . Thus, an arbitrary of the two items,  $a_3$  and  $a_4$ , say  $a_3$ , will be the second item that will be replaced with  $a'_3$  in the first step. In the second step only two items,  $a_2$  and  $a_4$ , are left, so both will be replaced with  $a'_2$  and  $a'_4$  in the second step. Nevertheless, let us illustrate our definitions also in the second step of the replacement procedure. The set  $\bar{A}_2$  consists of two remaining items:  $a_2$  and  $a_4$ . We have  $W_2 = \{((a_2, X_2), (a_4, X_2)), ((a_2, X_3), (a_4, X_3))\}$ . Naturally,  $s_{2,1}(a_2) = s_{2,1}(a_4) = s_{2,2}(a_2) = s_{2,2}(a_4) = 2$ .

We want now to derive bounds on the values  $s_{j,1}$  and  $s_{j,2}$ . The following inequalities hold:

1.  $s_{j,1} \geq \frac{2|W_j|}{2q_f - 2(j-1)}$  for each  $1 \leq j \leq q$ .

$W_j$  contains pairs of items belonging to  $\bar{A}_j$ .  $\bar{A}_1$  has  $2q_f$  items, and  $\bar{A}_{j+1}$  is obtained from  $\bar{A}_j$  by

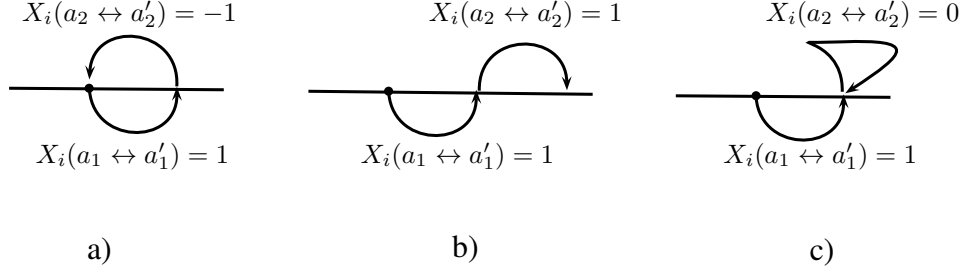


Figure 2: Figure illustrating that for  $X_i \in X_{\text{ex}}$ ,  $g_i(\{a_1, a_2\}, \{a'_1, a'_2\})$  is greater than  $(g_i(a_1, a'_1) + g_i(a_2, a'_2))$  if and only if  $((a_1, X_i), (a_2, X_i))$  is an annihilating pair. The figure presents 3 scenarios: a)  $((a_1, X_i), (a_2, X_i))$  is an annihilating pair. Both replacing  $a_1$  with  $a'_1$  and replacing  $a_2$  with  $a'_2$  moves us further from the target distribution for attribute  $X_i$  (the target distribution is marked as a black dot), thus  $g_i(a_1, a'_1) = -\frac{2}{k}$  and  $g_i(a_2, a'_2) = -\frac{2}{k}$ . However these changes annihilate, and  $g_i(\{a_1, a_2\}, \{a'_1, a'_2\}) = 0$ . b)  $g_i(a_1, a'_1) = -\frac{2}{k}$  and  $g_i(a_2, a'_2) = -\frac{2}{k}$ , but these changes do not annihilate, and thus:  $g_i(\{a_1, a_2\}, \{a'_1, a'_2\}) = -\frac{4}{k}$ . c)  $g_i(a_1, a'_1) = -\frac{2}{k}$  and  $g_i(a_2, a'_2) = 0$ , if at least one change does not move the solution against the target distribution, the changes do not annihilate, and  $g_i(\{a_1, a_2\}, \{a'_1, a'_2\}) = g_i(a_1, a'_1) + g_i(a_2, a'_2)$ .

removing two items. Consequently,  $\bar{A}_j$  has  $2q_f - 2(j-1)$  items, and thus,  $W_j$  contains pairs of  $2q_f - 2(j-1)$  different items. From the pigeonhole principle it follows that there exists an item that belongs to at least  $\frac{2|W_j|}{2q_f - 2(j-1)}$  pairs. Naturally, we also get the weaker constraint:  $s_{j,1} \geq \frac{|W_j|}{2q_f - 2(j-1)}$ .

2.  $|W_j| \geq |W_{j-1}| - 2s_{j-1,1}$  for each  $2 \leq j \leq q$ .

Each item in  $W_{j-1}$  belongs to at most  $s_{j-1,1}$  pairs (this follows from the definition of  $s_{j-1,1}$ ).  $W_j$  contains all pairs that  $W_{j-1}$  contained, except for the pairs involving  $a_{j-1,1}$ ,  $a_{j-2,2}$  (to obtain  $\bar{A}_j$ , we removed these two items from  $\bar{A}_{j-1}$ ). Consequently,  $W_j$  is obtained from  $W_{j-1}$  by removing at most  $2s_{j-1,1}$  pairs of items.

3.  $s_{j,2} \geq \frac{s_{j,1}}{2q_f - 2(j-1) - 1}$  for each  $1 \leq j \leq q$ .

In  $W_j$ , there are  $s_{j,1}$  pairs of items involving  $a_{j,1}$ . As we noted before,  $W_j$  contains pairs of  $2q_f - 2(j-1)$  different items. Thus, in  $W_j$ ,  $a_{j,1}$  is paired with at most  $2q_f - 2(j-1) - 1$  items. From the pigeonhole principle it follows that  $a_{j,1}$  must be paired with some item at least  $\frac{s_{j,1}}{2q_f - 2(j-1) - 1}$  times.

From Lemma 3 we get that:

$$\sum_{j=1}^q s_{j,2} \geq \frac{|W_1| \ln q}{4q}. \quad (8)$$

Before we proceed further let us make three observations regarding annihilating pairs. First, we note that for each  $X_i \in X_{\text{ex}}$ , and each  $a_x$  and  $a_y$ , if the value  $g_i(\{a_x, a_y\}, \{a'_x, a'_y\})$  is different from  $(g_i(a_x, a'_x) + g_i(a_y, a'_y))$  than it is greater from  $(g_i(a_x, a'_x) + g_i(a_y, a'_y))$  by  $\frac{4}{k}$ . We also note that  $g_i(\{a_x, a_y\}, \{a'_x, a'_y\})$  is greater than  $(g_i(a_x, a'_x) + g_i(a_y, a'_y))$  if and only if the changes  $X_i(a_x \leftrightarrow a'_x)$  and  $X_i(a_y \leftrightarrow a'_y)$  annihilate (this is illustrated in Figure 2). Further, we recall that the value  $s_{j,2}$  counts all attributes for which  $a_{j,1}$  and  $a_{j,2}$  constitute an annihilating pair. Thus, for each  $1 \leq j \leq q$ :

$$\sum_{i \in X_{\text{ex}}} g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) = \sum_{i \in X_{\text{ex}}} (g_i(a_{j,1}, a'_{j,1}) + g_i(a_{j,2}, a'_{j,2})) + s_{j,2} \frac{4}{k} \quad (9)$$

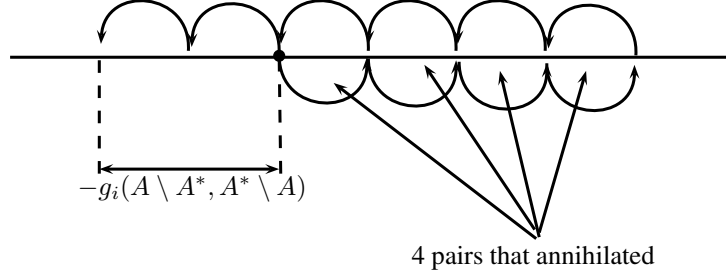


Figure 3: Figure illustrating the effect of replacing 10 items for an attribute  $X_i \in X_{\text{ex}}$ . Each replacement imposes a negative gain:  $g_i(a_j, a'_j) = -\frac{2}{k}$  for  $1 \leq j \leq 10$ . Thus,  $\sum_{a \in A \setminus A^*} g_i(a, a') = -\frac{20}{k}$ . In this example four pairs annihilated, and, consequently,  $g_i(A \setminus A^*, A^* \setminus A) = -\frac{4}{k}$ .

Our second observation is similar in spirit to the first one. We note that for each  $X_i \in X_{\text{ex}}$ :

$$g_i(A \setminus A^*, A^* \setminus A) - \sum_{a \in A \setminus A^*} g_i(a, a') = \text{the number of pairs that annihilated for } X_i \times \frac{4}{k}.$$

The above equality is illustrated in Figure 3. As a consequence, we get that:

$$\sum_{X_i \in X_{\text{ex}}} \left( g_i(A \setminus A^*, A^* \setminus A) - \sum_{a \in A \setminus A^*} g_i(a, a') \right) = \text{the number of pairs that annihilated} \times \frac{4}{k}.$$

We recall that after the replacement procedure  $A$  can differ from  $A^*$  with at most one element, hence, having distance to the optimal distribution at most equal to  $|X| \frac{2}{k}$ . Thus:

$$\sum_{X_i \in X_{\text{ex}}} \left( g_i(A \setminus A^*, A^* \setminus A) - \sum_{j=1}^q (g_i(a_{j,1}, a'_{j,1}) + g_i(a_{j,2}, a'_{j,2})) \right) \leq P \cdot \frac{4}{k} + |X| \frac{2}{k}. \quad (10)$$

Our third observation says that:

$$\sum_{X_i \in X_{\text{aex}} \setminus X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) - \sum_{X_i \in X_{\text{aex}} \setminus X_{\text{ex}}} \sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) \leq |X_{\text{aex}} \setminus X_{\text{ex}}|. \quad (11)$$

Where does Inequality 11 come from? Let us use the geometric interpretation, like the one from Figure 3. Let us consider an  $X_i$ ,  $X_i \in X_{\text{aex}}$ . For  $X_i$ ,  $A$  lies in a distance of  $\frac{2}{k}$  on the left or on the right from the target distribution. Without loss of generality, let us assume it lies on the right. Now, if  $g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) < 0$  then replacing  $(a_{j,1}, a_{j,2})$  with  $(a'_{j,1}, a'_{j,2})$  moves the current solution right. If  $g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) = \frac{2}{k}$ , then replacing  $(a_{j,1}, a_{j,2})$  with  $(a'_{j,1}, a'_{j,2})$  moves the current solution by  $\frac{2}{k}$  on left. If  $g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) = 0$ , then replacing  $(a_{j,1}, a_{j,2})$  with  $(a'_{j,1}, a'_{j,2})$  either does not move the solution or moves it by  $\frac{4}{k}$  on left.

Let us define  $y_i = g_i(A \setminus A^*, A^* \setminus A) - \sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\})$ . If the solution moves  $q$  times to the right, then the total gain  $-\sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\})$  will be maximized, achieving  $q \frac{4}{k}$ . In such case however, the value  $g_i(A \setminus A^*, A^* \setminus A)$  will be equal to  $-q \frac{4}{k}$ , and thus the value  $y_i$  will be equal to 0. After some consideration, the reader will see that the value  $y_i$  is maximized if the current solution moves  $\frac{q}{2}$  times right and  $\frac{q}{2}$  times left, each time by the value of  $\frac{4}{k}$ . This way, the moves to the right induce the total gain of  $\frac{q}{2} \cdot \frac{4}{k}$ , the moves to the left induce the zero gain, but as a consequence, the current solution for  $X_i$  does not change ( $g_i(A \setminus A^*, A^* \setminus A) = 0$ ). Thus, for each  $X_i \in X_{\text{aex}}$ ,  $y_i$  is upper bounded by  $\frac{q}{2} \cdot \frac{4}{k} \leq 1$ , which proves Inequality 11.

We can further proceed with the proof by observing that from the condition in the “while” loop we get that for each  $1 \leq j \leq q$ :

$$\begin{aligned} 0 &\geq \sum_i g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) \\ &\geq \sum_{i \in X_{\text{ex}}} g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) + \sum_{i \notin X_{\text{ex}}} g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) \end{aligned}$$

From Equality 9:

$$\geq \sum_{i \in X_{\text{ex}}} (g_i(a_{j,1}, a'_{j,1}) + g_i(a_{j,2}, a'_{j,2})) + s_{j,2} \frac{4}{k} + \sum_{i \notin X_{\text{ex}}} g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}).$$

Thus, we get:

$$- \sum_{i \in X_{\text{ex}}} (g_i(a_{j,1}, a'_{j,1}) + g_i(a_{j,2}, a'_{j,2})) - \frac{4}{k} s_j^2 > + \sum_{i \notin X_{\text{ex}}} g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}). \quad (12)$$

Next, we give the following sequence of inequalities:

$$\begin{aligned} g &= \sum_i g_i(A \setminus A^*, A^* \setminus A) \\ &= \sum_{X_i \in X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) + \sum_{X_i \in X_{\text{aex}} \setminus X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) + \sum_{X_i \notin X_{\text{aex}}} g_i(A \setminus A^*, A^* \setminus A) \end{aligned}$$

From Inequality 7, for all  $i \notin X_{\text{aex}}$ , we have  $g_i(A \setminus A^*, A^* \setminus A) \leq \sum_{a \in A \setminus A^*} g_i(a, a')$ . Since the set  $A \setminus A^*$  and  $\bigcup_{j=1}^q \{a_{j,1}, a_{j,2}\}$  can differ by at most one item (which induces distance  $\frac{2|X|}{k}$  to the optimal solution), we have that

$$\sum_{X_i \notin X_{\text{aex}}} g_i(A \setminus A^*, A^* \setminus A) \leq \sum_{X_i \notin X_{\text{aex}}} \sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) + \frac{2|X|}{k}.$$

And, as a consequence:

$$\begin{aligned} g &\leq \sum_{X_i \in X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) + \sum_{X_i \in X_{\text{aex}} \setminus X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) \\ &\quad + \sum_{X_i \notin X_{\text{aex}}} \sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) + \frac{2|X|}{k} \\ &\leq \sum_{X_i \in X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) + \sum_{X_i \in X_{\text{aex}} \setminus X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) \\ &\quad + \sum_{X_i \notin X_{\text{ex}}} \sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) - \sum_{X_i \in X_{\text{aex}} \setminus X_{\text{ex}}} \sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) + \frac{2|X|}{k}. \end{aligned}$$

From Inequality 11 we get:

$$g \leq \sum_{X_i \in X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) + \sum_{X_i \notin X_{\text{ex}}} \sum_{j=1}^q g_i(\{a_{j,1}, a_{j,2}\}, \{a'_{j,1}, a'_{j,2}\}) + \frac{2|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}|.$$

From Inequality 12:

$$g \leq \frac{2|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| + \sum_{X_i \in X_{\text{ex}}} g_i(A \setminus A^*, A^* \setminus A) - \sum_{X_i \in X_{\text{ex}}} \sum_{j=1}^q (g_i(a_{j,1}, a'_{j,1}) + g_i(a_{j,2}, a'_{j,2})) - \frac{4}{k} \sum_j s_{j,2}.$$

From Inequality 8:

$$g \leq \frac{2|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| - \frac{|W_1| \ln q}{4q} \cdot \frac{4}{k} + \sum_{i \in X_{\text{ex}}} \left( g_i(A \setminus A^*, A^* \setminus A) - \sum_{j=1}^q (g_i(a_{j,1}, a'_{j,1}) + g_i(a_{j,2}, a'_{j,2})) \right)$$

From Inequality 10:

$$g \leq \frac{4|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| - \frac{|W_1| \ln q}{kq} + P \frac{4}{k}.$$

As we noted before, from Lemma 2, we have that  $W_1 \geq \frac{P^2}{|X_{\text{ex}}|}$ . Thus:

$$g \leq \frac{4|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| + \frac{4}{k} \left( P - \frac{P^2 \ln q}{4|X_{\text{ex}}|q} \right).$$

Since  $q \leq \frac{k}{2}$ , and since the function  $\frac{\ln x}{x}$  is decreasing for  $x \geq 1$ :

$$g \leq \frac{4|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| + \frac{4}{k} \left( P - \frac{P^2 \ln(k/2)}{2|X_{\text{ex}}|k} \right)$$

The function  $f(P) = P - \frac{P^2 \ln(k/2)}{2|X_{\text{ex}}|k}$  takes its maximum for  $P = \frac{|X_{\text{ex}}|k}{\ln(k/2)}$ . Thus:

$$g \leq \frac{4|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| + \frac{4}{k} \cdot \frac{|X_{\text{ex}}|k}{2 \ln(k/2)} = \frac{4|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| + \frac{2|X_{\text{ex}}|}{\ln(k/2)}.$$

Since our local-search algorithm for  $\ell = 2$  also tries to perform local swaps on single items, we can repeat the analysis from the proof of Theorem 1. Thus, using Inequality 4 from there, we get that  $g \leq 2|X_{\text{ex}}|$ , and as a consequence:  $\left( \frac{1}{2} - \frac{1}{\ln(k/2)} \right) g \leq |X_{\text{ex}}| - \frac{2|X_{\text{ex}}|}{\ln(k/2)}$ .

For each attribute  $X_i \in X \setminus X_{\text{aex}}$  the distance from  $A$  and the target distribution is bounded by 2. For  $X_i \in X_{\text{aex}}$  this distance is bounded by  $\frac{2}{k}$ . Thus, we get that  $g \leq 2(|X| - |X_{\text{ex}}| - |X_{\text{aex}} \setminus X_{\text{ex}}|) + |X| \frac{2}{k}$ , and so:

$$\begin{aligned} g + \left( \frac{1}{2} - \frac{1}{\ln(k/2)} \right) g + \frac{1}{2} g &\leq \frac{4|X|}{k} + |X_{\text{aex}} \setminus X_{\text{ex}}| + \frac{2|X_{\text{ex}}|}{\ln(k/2)} \\ &\quad + |X_{\text{ex}}| - \frac{2|X_{\text{ex}}|}{\ln(k/2)} \\ &\quad + (|X| - |X_{\text{ex}}| - |X_{\text{aex}} \setminus X_{\text{ex}}|) + |X| \frac{2}{k} \\ &= |X| + \frac{6|X|}{k} \end{aligned}$$

Finally, we get:

$$g \leq \frac{\ln(k/2)}{2 \ln(k/2) - 1} \left( |X| + \frac{6|X|}{k} \right).$$

Which proves the thesis.  $\square$

Since a brute-force algorithm can be used to compute an optimal solution for small values of  $k$ , Theorem 2 implies that for every  $\epsilon > 0$  we can achieve an additive approximation of  $\frac{1}{2}(|X| + \epsilon)$ , that is we can guarantee that the solution returned by our algorithm will be at least 4 times better than a solution that is arbitrarily

bad on each attribute. A natural open question is whether the local search algorithm achieves even better approximation guarantees for larger values of  $\ell$ .

One may argue that the restriction to normal target distributions is a strong one. However, for a given vector of target distributions  $\pi$ , we can easily find a vector  $\pi_N$  of target normal distributions such that  $\|\pi, \pi_N\|_1 \leq \frac{2X}{k}$ . Thus, the results from Theorems 1 and 2 can be modified by providing approximation ratio worse by an additive value of  $\frac{2X}{k}$  but valid for arbitrary target distributions. Again, since an optimal solution can easily be computed for small values of  $k$ , we can get arbitrarily close to the approximation guarantees given by Theorems 1 and 2, even for non-normal target distributions.

Below we show a lower bound of  $\frac{2X}{7}$  for the approximation ratio of the local search algorithm from Figure 1 with  $\ell = 2$ .

**Example 3** Consider  $5p$  binary attributes  $X_1, \dots, X_{5p}$ ,  $6\ell$  and the set of distinct candidates  $C = \{a_1, \dots, a_\ell, a'_1, \dots, a'_\ell, b_1, \dots, b_\ell, b'_1, \dots, b'_\ell, c_1, \dots, c_\ell, c'_1, \dots, c'_\ell\}$  (in our database there exists a large number  $p$  of copies of each candidate from  $C$ ). For each  $i$ , we have:

	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
$a_i$	1	0	1	1	0	0	1
$a'_i$	0	1	0	0	1	1	1
$b_i$	0	0	0	0	0	0	0
$b'_i$	0	0	1	1	1	1	0
$c_i$	1	1	1	1	0	0	0
$c'_i$	1	1	0	0	1	1	0

We note that for each candidate the value of the attribute  $X_3$  is the same as of  $X_4$  and the value of the attribute  $X_5$  is the same as of  $X_6$ . For  $i \in \{1, 2, 3, 4, 5, 6\}$  let  $\pi_i^0 = \pi_i^1 = \frac{1}{2}$ , and let  $\pi_7^0 = 1 - \pi_7^1 = 1$ .

Let  $k = 4p$ . It can be easily checked that the set consisting of  $p$  copies of candidates  $b_i, b'_i, c_i, c'_i$  is a perfect committee. On the other hand, the set  $A$  consisting of  $2p$  copies of candidates  $a_i$  and  $a'_i$  is locally optimal. Indeed, replacing candidate  $a_i$  or  $a'_i$  with  $b_i$  or  $b'_i$  moves the solution closer to the target distribution on  $X_7$ , but the further from the target distribution on  $X_1$  or  $X_2$ . The same situation happens if we replace candidates  $a_i$  or  $a'_i$  with  $c_i$  or  $c'_i$ . If we replace two  $a$ -candidates with the pair consisting of one  $b$ -candidate ( $b_i$  or  $b'_i$ ) and one  $c$ -candidate ( $c_i$  or  $c'_i$ ), then such replacement will move the solution closer by  $\frac{4}{k}$  to the target distribution on  $X_7$ , but will move the solution further by  $\frac{2}{k}$  on two attributes from  $\{X_3, X_4, X_5, X_6\}$ .

Finally,  $\|\pi, r(A)\|_1 = 2p = \frac{2}{7}|X|$ .

### 6.3 Parameterized Complexity

In this section, we study the parameterized complexity of the problem of finding a perfect committee. We are specifically interested whether for some natural parameters there exist fixed parameter tractable (FPT) algorithms. We recall that the problem is FPT for a parameter  $P$  if its each instance  $I$  can be solved in time  $O(f(P) \cdot \text{poly}(|I|))$ .

From the point of view of parameterized complexity, FPT is seen as the class of easy problems. There is also a whole hierarchy of hardness classes,  $\text{FPT} \subseteq W[1] \subseteq W[2] \subseteq \dots$  (for details, we point the reader to appropriate overviews [9, 19, 11]).

Obviously, the problem admits an FPT algorithm for the parameter  $m$ . Now, we present a negative result for parameter  $k$  (committee size) and a positive result for the parameter  $p$  (number of attributes).

**Theorem 3** *The problem of deciding whether there exists a perfect committee is  $W[1]$ -hard for the parameter  $k$ , even for binary domains.*

*Proof.* By reduction from the  $W[1]$ -complete PERFECTCODE problem [5]. Let  $I$  be an instance of PERFECTCODE that consists of a graph  $G = (V, E)$  and a positive integer  $k$ . We ask whether there exists

$V' \subseteq V$  such that each vertex in  $V$  is adjacent to exactly one vertex from  $V'$  (by convention, a vertex is adjacent to itself). From  $I$  we construct the following instance  $I'$  of the perfect committee problem. For each  $v \in V$  there is a binary attribute  $X_v$  and a candidate  $c_v$ . For each  $u, v \in V$ ,  $X_v(c_u) = 1$  if and only if  $u$  and  $v$  are adjacent in  $G$ . We look for a committee of size  $k$ . For each  $v$ ,  $\pi_v^1 = 1 - \pi_v^0 = \frac{1}{k}$ . It is easy to see that perfect codes in  $I$  correspond to perfect committees in  $I'$ .  $\square$

**Theorem 4** *For binary domains, there is an FPT algorithm for the perfect committee problem for parameter  $p$ .*

*Proof.* Each item can be viewed as a vector of values indexed with the attributes; there are  $2^p$  such possible vectors:  $v_1, \dots, v_{2^p}$ . For each  $v_i$ , let  $a_i$  denote the number of items that correspond to  $v_i$ . Consider the following integer linear program, in which each variable  $b_i$  is the number of candidates corresponding to  $v_i$  in a perfect committee.

$$\begin{aligned}
& \text{minimize } \sum_{i=1}^{2^p} b_i \\
& \text{subject to:} \\
& \text{(a) : } b_i \geq 0 \quad 1 \leq i \leq 2^p \\
& \text{(b) : } b_i \leq a_i \quad 1 \leq i \leq 2^p \\
& \text{(c) : } \sum_{i=1}^{2^p} b_i = k \\
& \text{(d) : } \sum_{i: v_i[j]=1} b_i = \pi_i^1 \quad 1 \leq j \leq p
\end{aligned}$$

This linear program has  $2^p$  variables, thus, by the result of Lenstra [15, Section 5] it can be solved in FPT time for parameter  $p$ . This completes the proof.  $\square$

**Example 4** *Let  $p = 2$ ,  $k = 5$ , and let the candidate database  $C$  consists of 4 candidates with value vector  $v_1 = (0, 0)$ , 2 with value vector  $v_2 = (1, 0)$ , 2 candidates with value vector  $v_3 = (0, 1)$  and 2 candidates with value vector  $v_4 = (1, 1)$ . Let  $\pi = ((0.2, 0.8), (0.6, 0.4))$ . The integer linear program is*

$$\begin{aligned}
& \text{minimize } b_1 + b_2 + b_3 + b_4 \\
& \text{subject to:} \\
& \text{(a) : } b_i \geq 0 \quad 1 \leq i \leq 4 \\
& \text{(b) : } b_1 \leq 4; b_2 \leq 2; b_3 \leq 2; b_4 \leq 2 \\
& \text{(c) : } b_1 + b_2 + b_3 + b_4 = 5 \\
& \text{(d) : } b_1 + b_3 = 1; b_1 + b_2 = 3
\end{aligned}$$

*and a solution is  $(b_1 = 1, b_2 = 2, b_3 = 0, b_4 = 2)$ : a perfect committee is obtained by taking one candidate with value vector  $(0, 0)$ , two candidates with value vector  $(1, 0)$ , and two with value vector  $(1, 1)$ .*

*Now, consider the database  $C'$  consists of 5 candidates with value vector  $v_1 = (0, 0)$ , 2 with value vector  $v_2 = (1, 0)$ , 2 candidates with value vector  $v_3 = (0, 1)$  and 1 candidate with value vector  $v_4 = (1, 1)$ . Let  $\pi = ((0.2, 0.8), (0.6, 0.4))$ : then the corresponding constraints are inconsistent and there is no perfect committee.*

We conclude this Section by a short discussion. Finding an optimal committee is likely to be difficult if the candidate database  $C$  is large, and the number of attributes not small. Assume  $|C|$  is large compared to

the size of the domain  $\prod_{i=1}^P |D_i|$ , that each attribute value appears often enough in  $C$  and that there is no strong correlation between attributes in  $C$ : then, the larger  $|C|$ , the more likely  $C$  satisfies Full Supply, in which case finding an optimal committee is easy. The really difficult cases are when  $|C|$  is not significantly larger than the domain, or when  $C$  shows a high correlation between attributes.

## 7 Conclusion

We have defined, and studied, multi-attribute generalizations of a well-known apportionment method (Hamilton), albeit with motivations that go far beyond party-list elections (such as the selection of a common set of items). We have shown positive and negative results concerning the properties satisfied by these generalizations and their computation, but a lot remains to be done. Note that other largest remainder apportionment methods can be generalized in a similar way, but it is unclear how largest-average methods can be generalized.

## References

- [1] M. Balinski and P. Young. Criteria for proportional representation. *Operations Research*, 27(1):80–95, 1979.
- [2] M. Balinski and P. Young. *Fair Representation : Meeting the Ideal of One Man One Vote*. Brookings Institution Press, seconde edition, 2001.
- [3] N. Betzler, A. Slinko, and J. Uhlmann. On the computation of fully proportional representation. *JAIR*, 2013.
- [4] S. J. Brams. Computer-assisted constrained approval voting. *Interfaces*, 20(5):67–80, 1990.
- [5] M. Cesati. Perfect code is W[1]-complete. *Information Processing Letters*, 81(3):163–168, 2002.
- [6] B. Chamberlin and P. Courant. Representative deliberations and representative decisions: Proportional representation and the Borda rule. *American Political Science Review*, 77(3):718–733, 1983.
- [7] D. Cornaz, L. Galand, and O. Spanjaard. Bounded single-peaked width and proportional representation. In *ECAI*, pages 270–275, 2012.
- [8] N. Ding and F. Lin. On computing optimal strategies in open list proportional representation: The two parties case. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, pages 1419–1425, 2014.
- [9] R. Downey and M. Fellows. *Parameterized Complexity*. Springer-Verlag, 1999.
- [10] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. Properties of multiwinner voting rules. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2014)*, May 2014. Also presented in FIT-2013.
- [11] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer-Verlag, 2006.
- [12] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.
- [13] J. Lang and L. Xia. Voting over multiattribute domains. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 9. Cambridge University Press, 2015.
- [14] I. Lari, F. Ricca, and A. Scozzari. Bidimensional allocation of seats via zero-one matrices with given line sums. *Annals OR*, 215(1):165–181, 2014.



- [15] H. W. Lenstra. Integer programming with a fixed number of variables. *Mathematics of Operations Research*, 8(4):538–548, 1983.
- [16] T. Lu and C. Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, pages 280–286, 2011.
- [17] T. Lu and C. Boutilier. Multiwinner social choice with incomplete preferences. In *IJCAI*, 2013.
- [18] B. L. Monroe. Fully proportional representation. *American Political Science Review*, 89:925–940, 1995.
- [19] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford University Press, 2006.
- [20] R. Potthoff. Use of linear programming for constrained approval voting. *Interfaces*, 20(5):79–80, 1990.
- [21] A. Procaccia, J. Rosenschein, and A. Zohar. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30(3):353–362, 2008.
- [22] F. Pukelsheim, F. Ricca, B. Simeone, A. Scozzari, and P. Serafini. Network flow methods for electoral systems. *Networks*, 59(1):73–88, 2012.
- [23] P. Serafini and B. Simeone. Parametric maximum flow methods for minimax approximation of target quotas in biproportional apportionment. *Networks*, 59(2):191–208, 2012.
- [24] P. Skowron, P. Faliszewski, and A. Slinko. Achieving fully proportional representation: Approximability result. *Artificial Intelligence*, 222:67–103, 2015.
- [25] A. Straszak, M. Libura, J. Sikorski, and D. Wagner. Computer-assisted constrained approval voting. *Group Decision and Negotiation*, 2(4):375–385, 1993.